ERDC/CHL TR-12-19

Coastal and Hydraulics Laboratory

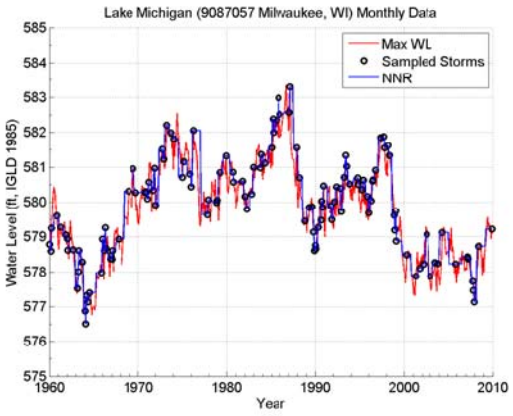
US Army Corps of Engineers®
Engineer Research and Development Center

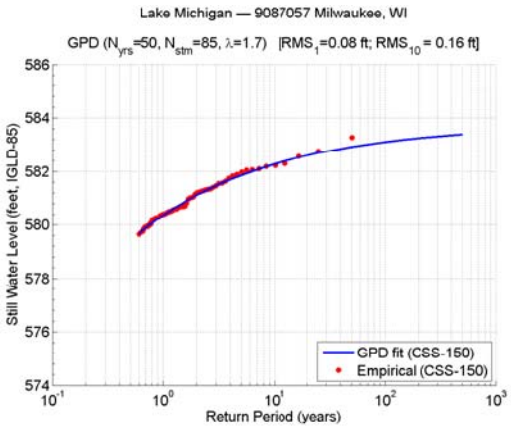
# Statistical Analysis and Storm Sampling Approach for Lakes Michigan and St. Clair

Great Lakes Coastal Flood Study, 2012 Federal Inter-Agency Initiative

Norberto C. Nadal-Caraballo, Jeffrey A. Melby, and Bruce A. Ebersole

September 2012



ERDC/CHL TR-12-19
September 2012

# Statistical Analysis and Storm Sampling Approach for Lakes Michigan and St. Clair

Great Lakes Coastal Flood Study, 2012 Federal Inter-Agency Initiative

Norberto C. Nadal-Caraballo, Jeffrey A. Melby,
and Bruce A. Ebersole

*Coastal and Hydraulics Laboratory*
*U.S. Army Engineer Research and Development Center*
*3909 Halls Ferry Road*
*Vicksburg, MS 39180-6199*

Final report



Prepared for U.S. Army Corps of Engineers
Detroit District
477 Michigan Ave.
Detroit, MI 48226

and

FEMA Region V
536 S. Clark St., 6th Floor
Chicago, IL 60605


FEMA

# Abstract

A new methodology is proposed for Great Lakes flood hazard mapping. The methodology includes a process for sampling and screening storm events and computing water level probabilities based on high-fidelity modeling of significant storm events. A technical analysis framework is provided to construct accurate extremal distributions of total water levels and to accurately estimate base flood elevations. High-resolution, high-fidelity modeling of all historical storm events is simply not feasible due to time, computational and funding constraints. Therefore, the recommended approach is to screen and sample historical events to select the minimum number of events required to accurately model the total water level extremal distributions. This study focused on evaluating the validity of the recommended statistical analysis and storm sampling approach, and determining the adequate storm sample size. For this purpose, several tasks were performed, including: computation of storm-surrogate waves and water levels; definition of full sample and composite storm sets; and evaluation of the statistical analysis approach for a record length of 50 years.

It was determined that the ideal number of events that should be sampled to accurately define water level extremal distributions is roughly 150 storms. Also, the prioritization of waves and water levels in storm sampling was extensively evaluated. It was found that approximately 25 percent to 30 percent of all selected storms were both highly ranked surge events and highly ranked wave events, thus minimizing the effects of different event prioritization ratios. Extreme water levels, corresponding to one percent and 0.2 percent annual chance of exceedance, had negligible variation regardless of the event prioritization ratio.

# Contents

# Figures and Tables

## Figures

## Tables

# Preface

The study summarized in this report was conducted at the request of the Detroit District (LRE), U.S. Army Corps of Engineers (USACE). Mary Weidel was LRE Program Manager and Greg Mausolf was the primary engineering point of contact at LRE. The study was funded by the Federal Emergency Management Agency (FEMA) through LRE and conducted at the Engineering Research and Development Center (ERDC), Coastal and Hydraulics Laboratory (CHL), Vicksburg, MS, during the period November 2010 – April 2012. The FEMA Lead was Ken Hinterlong, Chief, Risk Analysis Branch, Mitigation Division, FEMA Region V. Julie Tochor, Accenture, was the Program Management Lead for FEMA Region V.

This report was prepared by Dr. Norberto C. Nadal-Caraballo and Dr. Jeffrey A. Melby, Harbors, Entrances, and Structures (HES) Branch, CHL, and Bruce Ebersole, Chief, Flood and Storm Protection Division. Drs. Melby and Nadal-Caraballo were under the general supervision of Dr. Jackie Pettway, Chief, HES Branch, and Dr. Rose Kress, Chief, Navigation Division. Dr. William D. Martin was Director, CHL, and Jose Sanchez was Deputy Director, CHL.

COL Kevin J. Wilson was Commander and Executive Director of ERDC. Dr. Jeffery P. Holland was ERDC Director.

# Unit Conversion Factors

A sponsor requirement for this study was the use of English Customary units of measurement. Most measurements and calculations were done in SI units and then converted to English Customary. The following table can be used to convert back to SI units.

| Multiply | By | To Obtain |
|---|---|---|
| feet | 0.3048 | meters |
| cubic feet | 0.02831685 | cubic meters |
| pounds (force) | 4.448222 | Newtons |
| square feet | 0.09290304 | square meters |

# 1 Introduction

## 1.1 Overview

Developments in hydrodynamic modeling technology are changing flood hazard mapping. Only recently has it been conceivable to use high-fidelity planetary boundary layer models, storm surge models and wave generation/transformation models to hindcast all significant storms over many decades, and construct highly accurate extremal distributions of flood levels. This new technology represents breakthroughs in accuracy and in computational efficiency. In addition, our improved understanding of the physical processes and their variances has promoted more deliberate and defensible strategies for flood hazard mapping.

This report provides a detailed approach for flood hazard mapping in the Great Lakes. It reflects these recent advances in understanding the physical processes and our ability to model storm wave and surge generation and propagation at high resolution and high fidelity.

A new methodology is proposed for screening and sampling storm events and computing water level probabilities for the Great Lakes flood hazard mapping production. The methodology is based on data, process analysis, and trial methodology summarized by Melby et al. (2012). This method utilizes high-fidelity modeling, as described by Jensen et al. (2012), of a set of significant storm events to construct an accurate extremal distribution of total water levels. High resolution modeling of all historical storm events is simply not feasible due to time, computational and funding constraints. Thus, a major goal was to minimize the total number of storms modeled. Therefore, Melby et al. (2012) recommended an approach to screen and sample historical events to select the minimum number of events required to accurately model the total water level extremal distribution. The methodology is further developed below.

The study by Melby et al. (2012) showed that selecting between 15 and 20 unique storms for each of the nine NOAA water level gage locations was adequate to model the extremal distributions for the 27-year period from 1970 – 1997. This record was defined as the overlapping period of time for which continuous hourly measured water levels and 3-hour WIS hindcast waves were available, because these synoptic data were used as surrogates

to the high-fidelity modeling of storm flooding in the storm sampling process.

The one percent annual chance event refers to an event with an annual exceedance probability of one percent, meaning that there is a one percent chance of such event being equaled or exceeded within any given year. The one percent and 0.2 percent annual chance events are also denoted throughout this report as 100-year and 500-year events, respectively. In terms of its recurrence interval, a one percent annual chance event is referred to as a 100-year event, which is expected to be equaled or exceeded every 100 years, on average. The chance of a 100-year event occurring in any 100-year period is expected to be 100 percent (i.e., has an expected value of one). The recurrence interval of any event is calculated by dividing 100 by its percent annual chance. Thus, a 0.2 percent annual chance event corresponds to a 500-year return period.

Melby et al. (2012) recommended that a period of record of 50 years from 1960 to 2010 be used for selecting storms for the flood hazard mapping production, as this was shown to be the minimum record length to determine the proper shape of the extremal water level distribution. Although Melby et al. (2010) showed that the methodology yields accurate estimates of the one percent annual chance base flood elevation (BFE) and the 0.2 percent annual chance flood for a 27-year record, there were not sufficient data to verify the method for a 50-year record length. Therefore, the methodology is validated herein for the full 50-year period of record. The methodology is based on data from Lake Michigan and Lake St. Clair. This methodology should be generally applicable to the other Great Lakes. However, deviations are expected due to variations in physical processes.

This report, and the discussed methodology, resulted from a focus study conducted as part of the FEMA Region V/II flood hazard mapping project. The methodology represents a consensus of the FEMA Region V/II technology evaluation team consisting of members from the USACE Detroit District, USACE ERDC, Accenture, FEMA, and joint ventures RAAMP and STARR. These joint ventures consist of prominent U.S. engineering consulting firms with a long history of flood hazard mapping experience. This consensus was reached after a number of meetings and thorough technical review by STARR and RAAMP. The STARR and RAAMP reviews included independent focus studies at other lakes.

# 2 Data Summary

## 2.1 Datum

All water levels are referenced to the International Great Lakes Datum 1985 (IGLD85). A description of the datum can be found on the USACE Detroit District web site[1]. The Low Water Datum (LWD) defined for each Lake is as listed in Table 1.

Table 1. Elevations of low water datum on the Great Lakes.

| Lake | Low Water Datum Elevation | |
|---|---|---|
| | Feet Above IGLD85 | Feet Above NGVD88 (approximate) |
| Lake Superior | 601.1 | 601 |
| Lake Michigan | 577.5 | 578 |
| Lake Huron | 577.5 | 578 |
| Lake St. Clair | 572.3 | 572 |
| Lake Erie | 569.2 | 569 |
| Lake Ontario | 243.3 | 243 |

## 2.2 Water level data

Water level measurements from nine long-term National Oceanic and Atmospheric Administration (NOAA) stations were used for this study. Data collected since 1970 were recorded at hourly intervals. Pre-1970 data are only readily available as monthly averages and maxima. The term *monthly maxima* is herein defined as the maximum hourly value for each month. For this study, monthly data were used from 1960 to 1969. Water level values were adjusted using the latest Basis of Comparison (BOC) levels, dated February 2011, which were developed by the International Upper Great Lakes Study and provided by the USACE Detroit District. BOC adjustments are required to reflect present day hydraulic conditions (e.g., outflow regulation, dredging, and diversions). NOAA Station names, ID, and locations are shown in Table 2. The station locations are shown in Figure 1 and the data are analyzed in the following sections.

[1] http://www.lre.usace.army.mil/greatlakes/hh/newsandinformation/igldatum1985/

Table 2. Location of NOAA water level gages for Lake Michigan.

| Station | Station ID | Latitude | Longitude | Hourly Records |
|---|---|---|---|---|
| Mackinaw City | 9075080 | 45.777 N | 84.725 W | 1970 – 2010 |
| Ludington, MI | 9087023 | 43.947 N | 86.442 W | 1970 – 2010 |
| Holland, MI | 9087031 | 42.767 N | 86.200 W | 1970 – 2010 |
| Calumet Harbor, IL | 9087044 | 41.728 N | 87.538 W | 1970 – 2010 |
| Milwaukee, WI | 9087057 | 43.002 N | 87.887 W | 1970 – 2010 |
| Kewaunee, WI | 9087068 | 44.463 N | 87.500 W | 1973 – 2010 |
| Sturgeon Bay, WI | 9087072 | 44.795 N | 87.313 W | 1970 – 2010 |
| Green Bay, WI | 9087079 | 44.540 N | 88.007 W | 1970 – 2010 |
| Port Inland, MI | 9087096 | 45.968 N | 85.870 W | 1970 – 2010 |

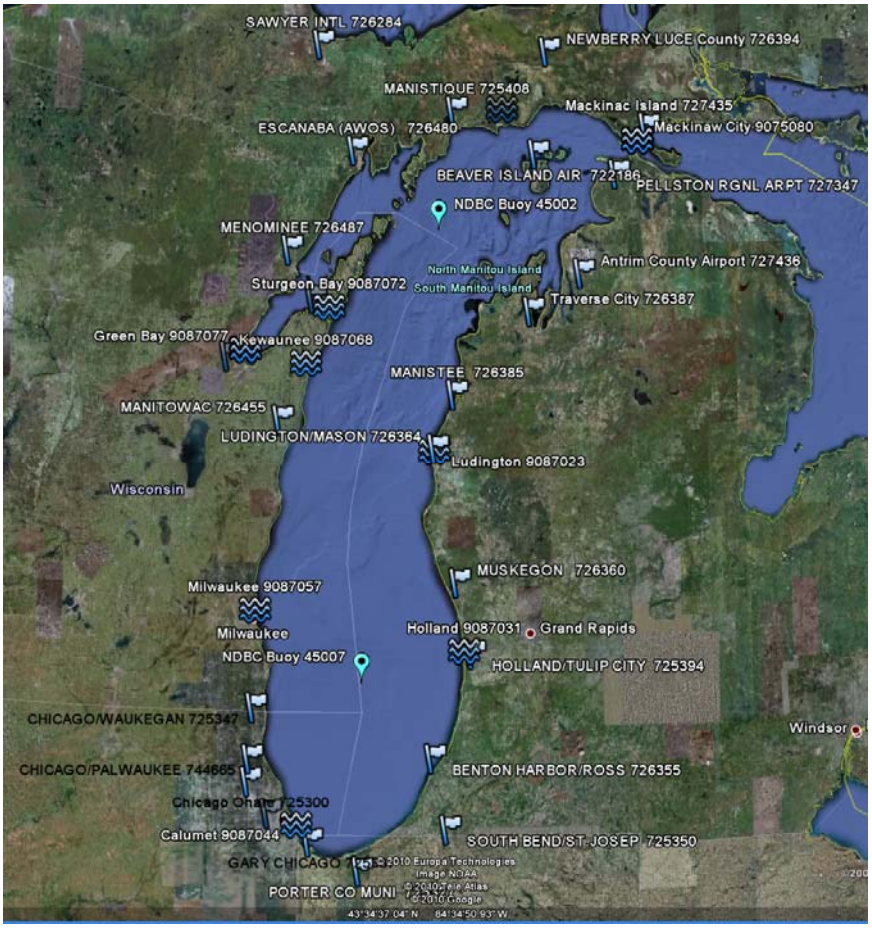



Figure 1. Locations of NOAA water level gages (three wavy lines), wave rider buoys (inverted teardrop) and meteorological (met) stations (flag) on Lake Michigan.

## 2.3 Meteorological data

The meteorological data used in this study were acquired from the NOAA National Climatic Data Center (NCDC) land-based stations, and are summarized in Table 3.

Table 3. Location of WBAN land-based meteorological stations for Lake Michigan.

| Station | Station ID | Latitude | Longitude | Hourly Records |
|---|---|---|---|---|
| Chicago/O'Hare | 725300/94846 | 41.986 N | 87.914 W | 1956 – 2010 |
| Chicago/Midway | 725340/14819 | 41.783 N | 87.750 W | 1948 – 2010 |
| Traverse City/Cherry | 726387/14850 | 44.741 N | 85.583 W | 1949 – 2010 |
| Milwaukee/Mitchell | 726400/14839 | 42.947 N | 87.897 W | 1950 – 2010 |
| Green Bay/Straubel | 726450/14898 | 44.513 N | 88.120 W | 1956 – 2010 |
| Mackinac Island | 727435/54820 | 45.865 N | W | 1956 – 2010 |

# 3 Water Level – Probability Distributions

As discussed in Melby et al. (2012), there are two basic methods to sample extreme values from a continuous record: Annual Maximum Series (AMS) and Partial Duration Series (PDS). However, there are significant pitfalls associated with the use of AMS as discussed in Melby et al. (2012). Valid extreme storms are discarded when more than one extreme storm occurs in a given year and minor events are included during mild years. This produces the same effect as having too short of a record length by steepening the extremal distribution of total water level and introducing potentially large errors in BFE estimation.

Melby et al. (2012) recommended using a PDS developed from the peaks-over-threshold (POT) sampling method. According to extreme value theory, the PDS determined from the POT method should conform to the Generalized Pareto Distribution (GPD). Melby et al. (2012) showed that PDS/GPD fits of extremes are considerably more accurate than those derived using the AMS/Generalized Extreme Value (GEV) distribution approach.

The typical criticism of the POT method is that the threshold is arbitrary. However, adoption of recently developed methods makes it deliberate and repeatable. The approach proposed here involves the following steps:

1. Use an initial, low threshold to ensure all significant extreme events are being identified and sampled.
2. Use inter-event time of approximately 48 hours to reject duplicate storms and make sure that all remaining storms are independent and identically distributed (i.i.d.).
3. Fit GPD to the extremal events using the Maximum Likelihood Method (MLM).
4. After a preliminary GPD fitting, higher thresholds are evaluated to optimize the GPD fit. This optimization approach includes estimation of root-mean-square (RMS) deviation, and use of quantile-quantile (Q-Q) plots.

## 3.1 Still water level vs. total water level

The analysis of extremal events discussed throughout this report was performed with two distinct water level data sets: still water level (SWL), and total water level (TWL). The term SWL refers to lake-wide water level plus storm surge elevation. When wave runup, including wave setup, is superimposed on SWL it is then known as TWL. Herein, the methods used to calculate wave runup also intrinsically include wave setup. Wave runup models and empirically-derived equations typically include wave setup, but in those cases where they do not, a method to separately compute wave setup must be employed.

The TWL, along with additional considerations such as wave overtopping and erosion, is the reference elevation that is used to establish (BFEs). Both SWL and TWL elevations are illustrated in Figure 2, and are discussed in Chapters 10 and 11.

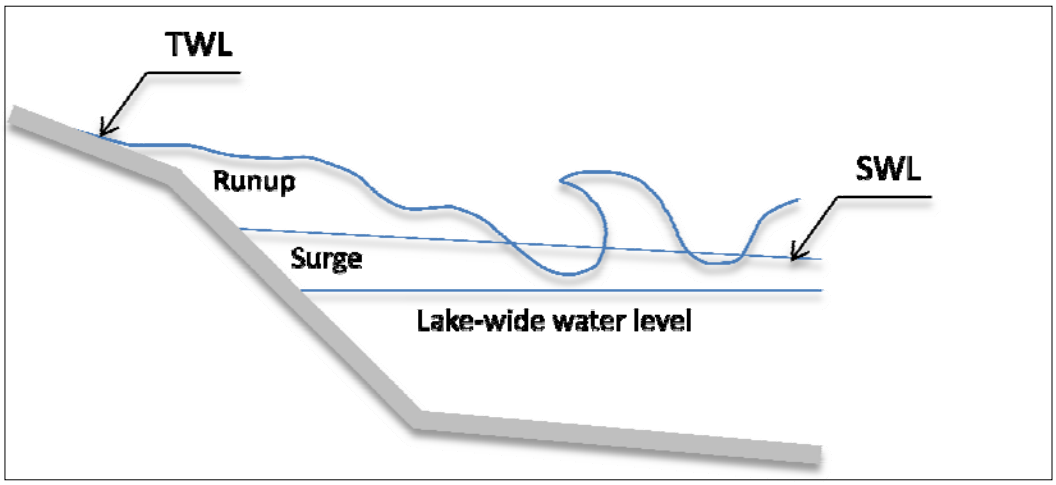



Figure 2. Still water and total water levels sketch.

When applying the POT method to generate the PDS of water levels, it is important to recognize which water level is being evaluated (i.e., SWL or TWL) and the type of storm data set that is sought. Full storm sets (FSS) are defined as sets of storms where all events are sampled from data representing the same spatial location. In other words, all parameters (waves and water levels) are computed at roughly the same spatial location. The FSS water level probability distribution represents the true distribution, or at least as close as one can get to the true distribution. On the other hand, composite storm sets (CSS) are a composite of extreme storms sampled from different gage locations spread across an area of interest (e.g., Lake Michigan). The main idea behind CSS is to sample

enough storms, and the right storms, to generate distributions that adequately match the FSS distributions and, thus, the true distributions all around the entire periphery of the lake.

FSS and CSS were generated for both SWL and TWL, as discussed in the following chapters. The length of the data records used varied according to the type of water level and the parameters involved. In the case of Lake Michigan, for example, SWL-FSS distributions can be constructed for the entire record lengths where water level measurements are available (approximately 37 to 107 years, depending on the gage). These would not be necessarily comparable to each other because of the different record lengths.

TWL-FSS distributions, however, have certain record length limitations. To construct a TWL-FSS, it is necessary to obtain wave measurements, wave hindcasts, or wind data to compute wave heights for the entire record length. For Lake Michigan, wave measurements and wave hindcasts are very limited (e.g., no wave data for Green Bay, for example). Wave measurements, wave hindcasts, and surrogate waves generated from measured winds are utilized to estimate waves. Reasonably accurate wind data are only available from 1960 to the present; therefore, TWL-FSS and TWL-CSS distributions can only be generated for this 50-year period.

For each of the water level sets (i.e., SWL-FSS, SWL-CSS, TWL-FSS, and TWL-CSS), two variants were considered and generated: (1) group of storms occurring throughout the year, and (2) group of storms exclusive of convective (summer) events. This was done for comparison purposes. Since fast-moving summer events (frontal systems and squall lines) cannot be properly modeled with currently available wind forcing data (i.e., insufficient temporal and spatial resolution), it is necessary to assess how their exclusion affects the probability distributions.

Also, note that SWL-CSS and TWL-CSS depend on the record length for which storms will be modeled. Based on the work of Melby et al. (2012), it was determined that storms would be modeled from 1960 on, so all composite storm sets are also limited to a 50-year period. The significance of storm data record length is discussed below.

## 3.2 Significance of record length

The exceedance probability of the most extreme storms is usually uncertain because it is a function of the record length, and to a lesser degree, of the number of samples. These extreme storms anchor the low-frequency tail of the distributions, so it is critical that they have the correct probability. A simple analysis of record length done by Melby et al. (2012) showed that the probability of the most extreme events was reasonably accurate for record lengths of 50 years or more. For decreasing record lengths, the extremes generally become more outliers and the return period distribution steepens. For very short record lengths (e.g., 30 years or less years) the return period distribution becomes unnaturally steep.

Melby et al. (2012) suggested that 50 year record lengths were the minimum length required to yield acceptable accuracy of distribution shapes for total water level. Storm surge distributions vary considerably as a function of record length, but water level distributions are asymptotic in the extreme tail. Therefore, 50 years should be generally adequate for total water level. For the present study, the 1960-2010 time period is recommended.

# 4 Storm Surrogates – Wind Data

Detailed hydrodynamic modeling of storms is expensive and time consuming and so should be limited to only the storms required to accurately describe the extremal total water level distribution. The most extreme storms are easily identified but the intermediate events must be selected judiciously. Selecting storms is complicated by the fact that measured or hindcast wave and water level data are not available for the full 50-year history from 1960-2010 to be modeled as part of the Great Lakes study, and measurements availability varies by regions. However, measured and hindcast winds are available for Lake Michigan for this 50-year period.

Measured or modeled winds can be used in different ways as part of a storm selection process. In this study measured winds were used to estimate storm surge and wave heights with relatively simpler *surrogate* calculation methods and were then compared against their measured or more rigorously hindcast counterparts. It is important to note that the surrogate calculation techniques are only used to facilitate selection of storms, which are to be simulated with computationally-intensive, high-resolution high-fidelity modeling. The following is a description of the process involved.

To define FSS and CSS and to conduct the work of this focus study, water level, surge elevation and wave height time series were sought at each of the nine NOAA water level gage locations used for the extremal analysis of storms in Lake Michigan. The main objective was to use wind data to estimate surge and waves to assemble synoptic data sets for the entire 50-year period.

Wind data spanning this period of time were available from six Weather Bureau-Army-Navy (WBAN) land-based methodological stations, as listed in Table 3. However, the spatial locations of these six meteorological stations differed from the locations of the nine NOAA gages (Table 2) where the synoptic wave and water level data were sought. It was thus necessary to spatially interpolate/extrapolate the wind data (i.e., magnitude and direction), as well as atmospheric pressure data, from their location of origin (Table 3) to the NOAA gage locations.

This interpolation/extrapolation was a computationally-intensive process because it was performed for the entire wind data record length (1960-2010) at 1-hour intervals. All spatial interpolation/extrapolation processes, as well as surge and waves computations were custom coded for this study in a series of Matlab® scripts.

As part of the storm surrogate analysis, surge and waves were estimated at 1-hour intervals continuously from 1960-2010. Thus, both non-convective (winter) and convective (summer) events were included, although it is unclear how well wind fields associated with convective events were represented by the interpolation/extrapolation techniques. The interpolation/extrapolation techniques are better suited for larger, more well-defined non-convective low pressure systems.

Once the wind data were spatially transformed, surge and wave height were estimated to assemble the synoptic storm data sets. The estimation of surge and waves are explained in Chapters 5 and 6, respectively.

## 4.1 Spatial transformation of wind data

The general steps followed as part of the wind data preprocessing are:

5. Acquired wind data from meteorological stations. These data included, at least: location, wind magnitude, wind direction, atmospheric pressure, information regarding the elevation at which the data were measured (any vertical datum), and how data were recorded (e.g., 20-minute averages).
6. Identified locations and coordinates where synoptic data sets are required (e.g., at NOAA water level gages).
7. Used interpolation/extrapolation to transform data from original locations (e.g., meteorological stations) to the end locations (e.g., at NOAA water level gages).

The wind velocity magnitude and direction were converted to vector form ($u$,$v$). Then the Matlab® function "gridfit.m" (Mathworks 2010) was used for the interpolation and extrapolation processes. This function employs a triangular scheme for data interpolation. For data extrapolation the "gridfit.m" function uses a spring-regularizer model that tends to drag the surface toward the mean of all the data. This avoids the large extrapolation errors that are typically associated with the use of gradient-based extrapolation schemes.

## 4.2 Meteorological data corrections

Meteorological parameters, such as wind speed, may require a series of adjustments before being used as part of any surge or wave height computation methodology. The adjustment factors applied in this study to the spatially transformed wind data are discussed below.

### 4.2.1 Equivalent neutral wind speed adjustment

This adjustment relates to the elevation at which wind speed measurements are taken. Most surge and wave height computation methodologies, including the Automated Coastal Engineering System (ACES; Leenkenecht et al. 1992) and Coastal Engineering Manual (CEM; USACE 2002) methodologies, utilize wind measurements recorded at 10-meter (32.8 ft) elevation, following the international standard reference height for winds. In those cases where measurements are recorded at elevations different than this reference elevation, the following equivalent neutral wind speed conversion must be performed (Schwab and Morton, 1984; CEM Part II, Equation II-2-9):

$$U_{10} = U_Z \left( \frac{10}{z} \right)^{1/7} \tag{1}$$

where

$U_{10}$ = wind speed observed at 10-meter reference elevation (m/sec)
$U_Z$ = wind speed at elevation $z$ different from reference elevation (m/sec)

### 4.2.2 Overland to overlake wind adjustment

Virtually all meteorological measurements near the Great Lakes are recorded at overland stations. Several methods exist for determining overlake wind speed as a function of the measured overland wind speed. The method used in this study was evaluated by Schwab and Morton (1984). Originally, Resio and Vincent (1977) used theoretical results derived from Cardone (1969) to develop curves relating overland to overlake wind speeds. Later, Schwab (1978) proposed the following equation as an approximation to the Cardone curves:

$$U_W = U_L\left(1.2+\frac{1.85}{U_L}\right)\left\{1-\frac{\Delta T}{|\Delta T|}\left(\frac{|\Delta T|}{1920}\right)^{1/3}\right\} \quad (2)$$

where

$U_W$ = overlake wind speed (m/sec)
$U_L$ = overland wind speed (m/sec)
$\Delta T$ = difference between overland air and water temperatures (°C)

Schwab's (1978) method was developed for overwater wind speed at an elevation of 20 m, therefore, these results must me multiplied by $(10/20)^{1/7}$ (equivalent neutral wind adjustment factor) to convert them to equivalent wind speeds at a 10-m elevation.

### 4.2.3 Wind speed averaging duration adjustment

The ACES/CEM wave growth methodology requires wind speeds to be 1-hour averages. If the measured or modeled wind data are averaged at different time intervals (e.g., 10-minute averages), the application of the following adjustment factor is required to correct wind speeds (CEM Part II, Figure II-2-1; ACES, Equation 21):

$$\frac{U_i}{U_{3600}} = 1.277+0.296\,tanh\left(0.9log\frac{45}{t_i}\right) \quad (3)$$

for $(1 < t_i < 3600\ sec)$

$$\frac{U_i}{U_{3600}} = -0.15\,log(t_i)+1.5334 \quad (4)$$

for $(3600 < t_i < 36000\ sec)$

where

$U_{3600}$ = 1-hour wind speed (m/sec, fps)
$U_i$ = wind speed at the desired duration of interest $t_i$ (m/sec, fps)
$t_i$ = duration of wind speed averaging (sec

# 5 Storm Surrogates – Storm Surge

Storm surge is defined in this study as a deviation of the measured water level time series from a moving average of the same time series. This moving average was generated by using a Gaussian smoothing algorithm with a 30-day time window. For the storm surrogate analysis, however, the estimation of storm surge was based on Shore Protection Manual (SPM; USACE 1984) guidelines, and Dean and Dalrymple (1990; 2001). The SPM, and Dean and Dalrymple discuss several components of, or contributors to, storm surge, including: barometric pressure, wind stress, and Coriolis force. For the purposes of this study, the storm surge was computed from steady-state equations for wind-shear stress and barometric pressure components only.

## 5.1 Storm surge components

The following are the steady-state equations for each component:

### 5.1.1 Wind-shear stress

The wind-shear stress component is due to the frictional drag generated by the wind blowing over the water surface. The empirical equation for estimating the wind-shear stress component is (Dean and Dalrymple, 1990):

$$\tau_w = \rho k W^2 \tag{5}$$

where

$\tau_w$ = wind-shear stress (N/m$^2$, lbf/ft$^2$)
$\rho$ = mass density of water (kg/m$^3$, lb/ft$^3$)
$W$ = wind speed at reference elevation of 10 m (32.8 ft)
$k$ = friction factor (non-dimensional)

The friction factor typically ranges from 1.2 x $10^{-6}$ to 3.4 x $10^{-6}$. Based on study results by Van Dorn (1953), the friction factor can be determined as (Dean and Dalrymple, 1990):

$$k = \begin{cases} 1.2\times10^{-6}, |W| \ll W_c \\ 1.2\times10^{-6} + 2.25\times10^{-6}\left(1-\frac{W_c}{|W|}\right)^2, |W| > W_c \end{cases} \quad (6)$$

where $W_c$ = 5.6 m/sec.

The magnitude of the wind-shear stress component, η, is obtained once the following implicit equation is solved (Dean and Dalrymple, 1990):

$$\frac{x}{\ell} = \left(1-\frac{h+\eta}{h_0}\right) - A\ln\left(\frac{\frac{h+\eta}{h_0}-A}{1-A}\right) \quad (7)$$

where

$\eta$ = surge elevation above mean water level due to wind-shear stress (m, ft)
$A$ = wind-shear to hydrostatic force ratio (non-dimensional)
$h$ = linearly varying water depth (m, ft)
$h_0$ = constant deep-water depth (m, ft)
$x$ = distance perpendicular to shoreline; x = 0 at shoreline (m, ft)
$\ell$ = continental shelf width (m, ft)

The water depth, $h$, as used in this steady-state equation is assumed to vary linearly from zero at the shoreline to a maximum value equal to the deep-water depth, $h_0$, at $x = \ell$.

The wind-shear stress is included in the parameter $A$, which is defined as:

$$A = \frac{n\tau_w \ell}{\rho g h_0^2} \quad (8)$$

where $n$ is a factor introduced to account for bottom-shear stress, which usually ranges from 1.15 to 1.30 (Dean and Dalrymple, 2001). The surrogate analysis performed as part of this study used $n$ = 1.25.

### 5.1.2 Barometric pressure

The barometric pressure component is due primarily to the pressure reduction at the center of the storm, where water is drawn up. The barometric response can be approximated as (Dean and Dalrymple, 2001):

$$\eta_b = \frac{\Delta p}{\gamma} \tag{9}$$

where

$\eta_b$ = surge component due to barometric pressure (m)
$\Delta p$ = barometric pressure difference (Pa)
$\gamma$ = specific weight of water (N/m$^3$)

The following rule of thumb is suggested by Dean and Dalrymple (2001): $\eta_b = 1.04\,\Delta p$, where $\eta_b$ is in centimeters and $\Delta p$ in millibars. The pressure difference was computed, as part of the calibration process, using different time lags ranging from 6 to 96 hours. A time lag of 48 hours was finally adopted in this study. For example, $\Delta p$ at $t$ = 0 hr was computed as the difference between $p$ at $t$ = 0 hr and $p$ at $t$ = -48 hr.

### 5.1.3 Coriolis force

The governing equation for the Coriolis force component is (Dean and Dalrymple, 2001):

$$\frac{\partial \eta_c}{\partial x} = \frac{\mathcal{F} V}{g} \tag{10}$$

where

$x$ = distance perpendicular to shoreline (m, ft)
$\mathcal{F}$ = is the Coriolis parameter (1/rads)
$V$ = magnitude of depth-averaged current (m/sec, fps)
$g$ = gravitational acceleration (m/sec$^2$, ft/sec$^2$)

The Coriolis force was not included as part of the storm surge computations due to its complexity and dynamic nature.

## 5.2 Storm surge results

As previously noted, for this study the storm surge was computed from steady-state equations with wind-shear stress and barometric pressure components. The total storm surge elevation, as estimated in this study, is the sum of the surge elevation above mean water level due to wind-shear stress ($\eta$) and the surge elevation due to barometric pressure ($\eta_b$). In other words, the total surge elevation is computed as $\eta_{tot} = \eta + \eta_b$.

Although significant effort was put behind the calibration process, the steady-state surrogate storm surge approach used in this study did not yield the desired results. The use of numerical models that take into account the dynamic effects of storm surge would be required to improve the results.

Figures 3 and 4 show comparisons between "measured" storm surge (from Gaussian smoothing) and estimated storm surge (from surrogate analysis) for the top 20 and top 100 storms for the Ludington gage, respectively. Low correlation between measured storm surge and storm surge calculated using the surrogate approach can be seen in both of these plots.

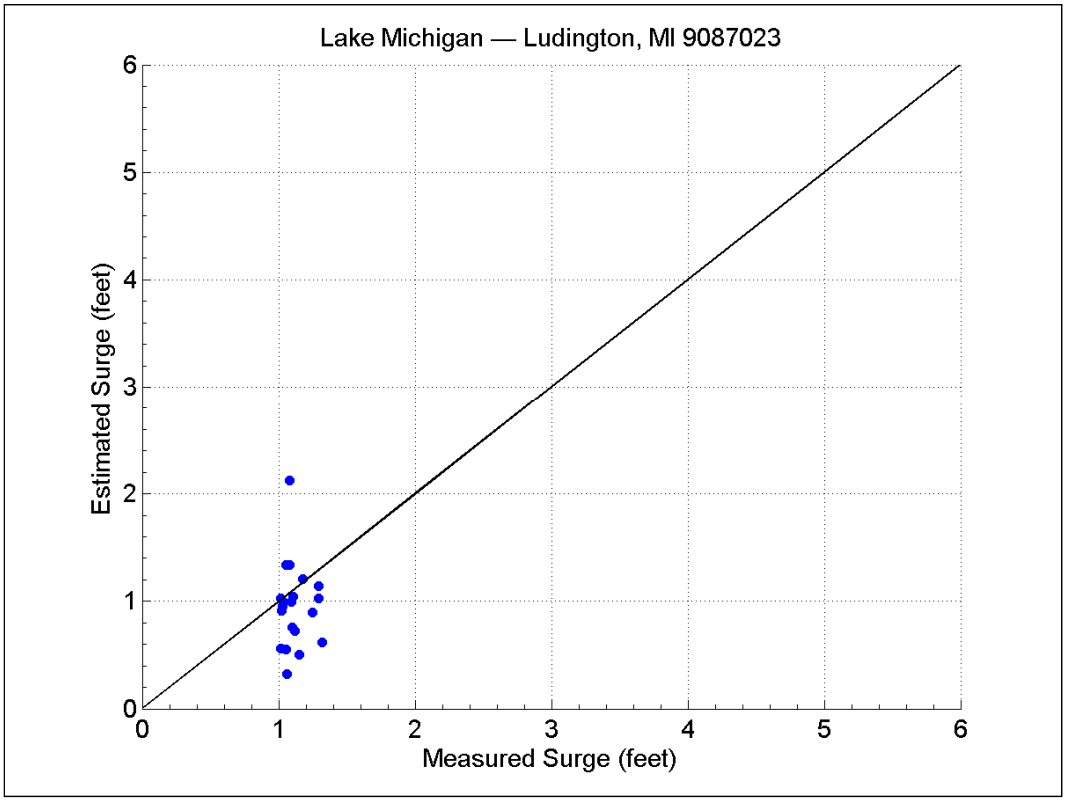



Figure 3. Top 20 surge events for Ludington, MI gage.

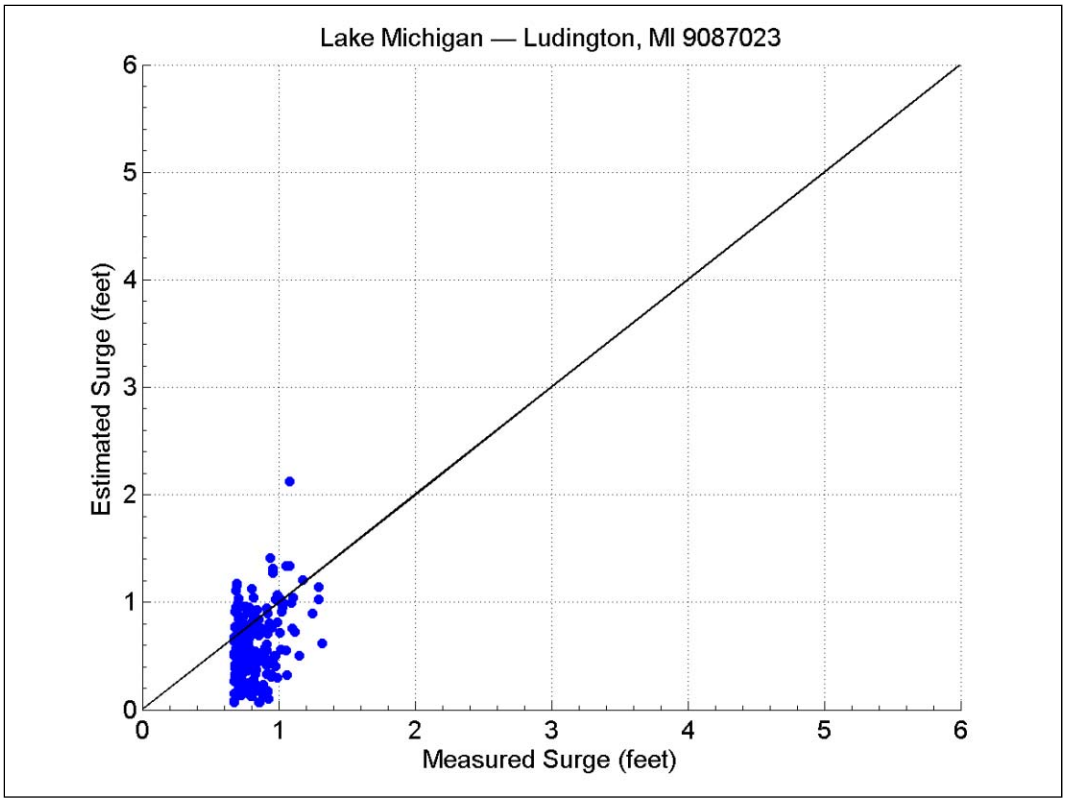



Figure 4. Top 100 surge events for Ludington, MI gage.

Figures 5 and 6 compare measured (FSS) data versus the estimated values for the top 20 and top 100 storms for the Milwaukee gage, respectively. Similar to the Ludington case, wide spreading and low correlation can be seen in both of these plots.

Figures 7 and 8 compare the return period plots of measured storm surge data (FSS) to steady state surge surrogate estimates for the Ludington gage. Similarly, Figures 9 and 10 compare the return period plots of measured data (FSS) to steady state surrogate estimates for the Milwaukee gage.

In the case of Ludington, the differences between the GPD fits shown on Figures 6 and 7 are evident. For Milwaukee, although the shapes of the GPD fits seem similar, the magnitudes of the surge events were still not estimated with adequate accuracy. The storms selected through POT analysis to construct the return period plot shown in Figure 10 were very different from the storms that correspond to the actual FSS (Figure 9). Therefore, the similarity between GPD fits was coincidental.

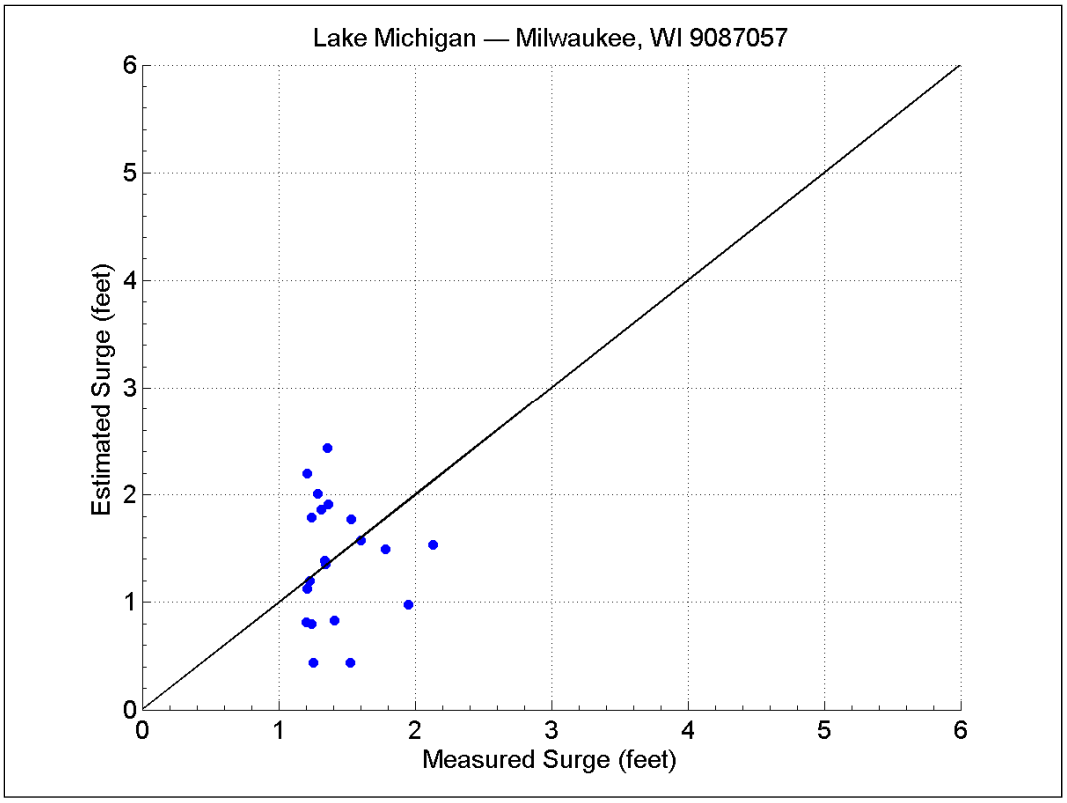



Figure 5. Top 20 surge events for Milwaukee, WI gage.

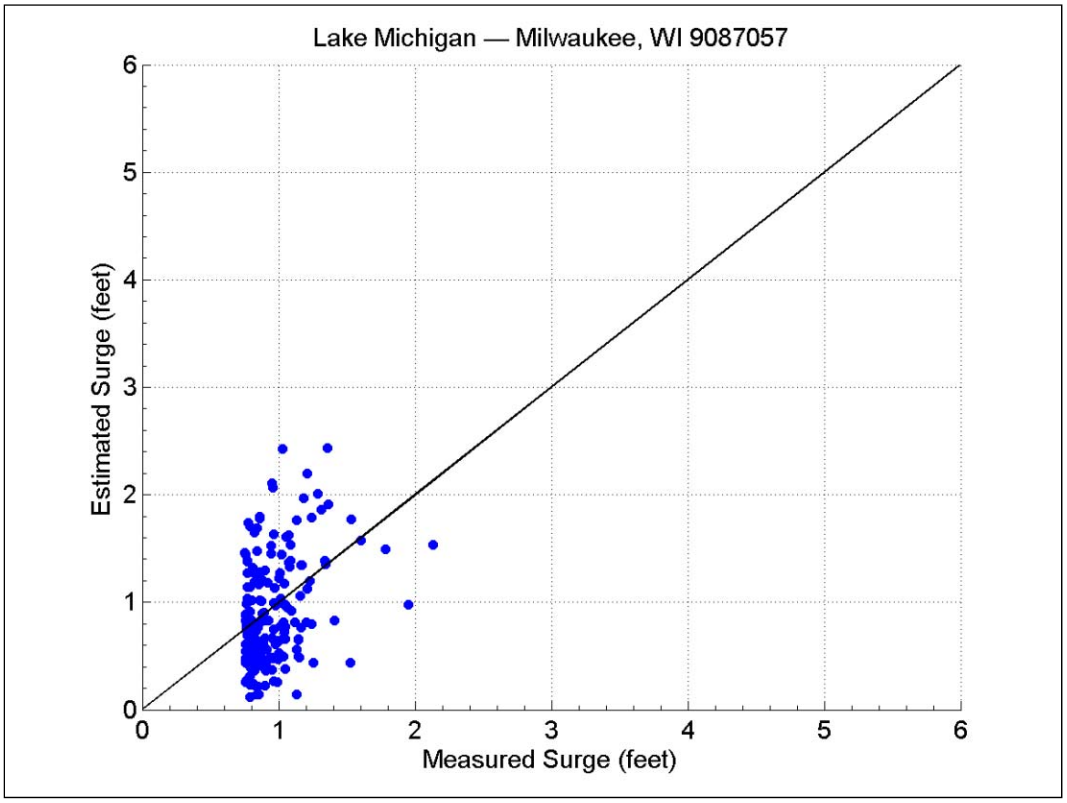



Figure 6. Top 100 surge events for Milwaukee, WI gage.

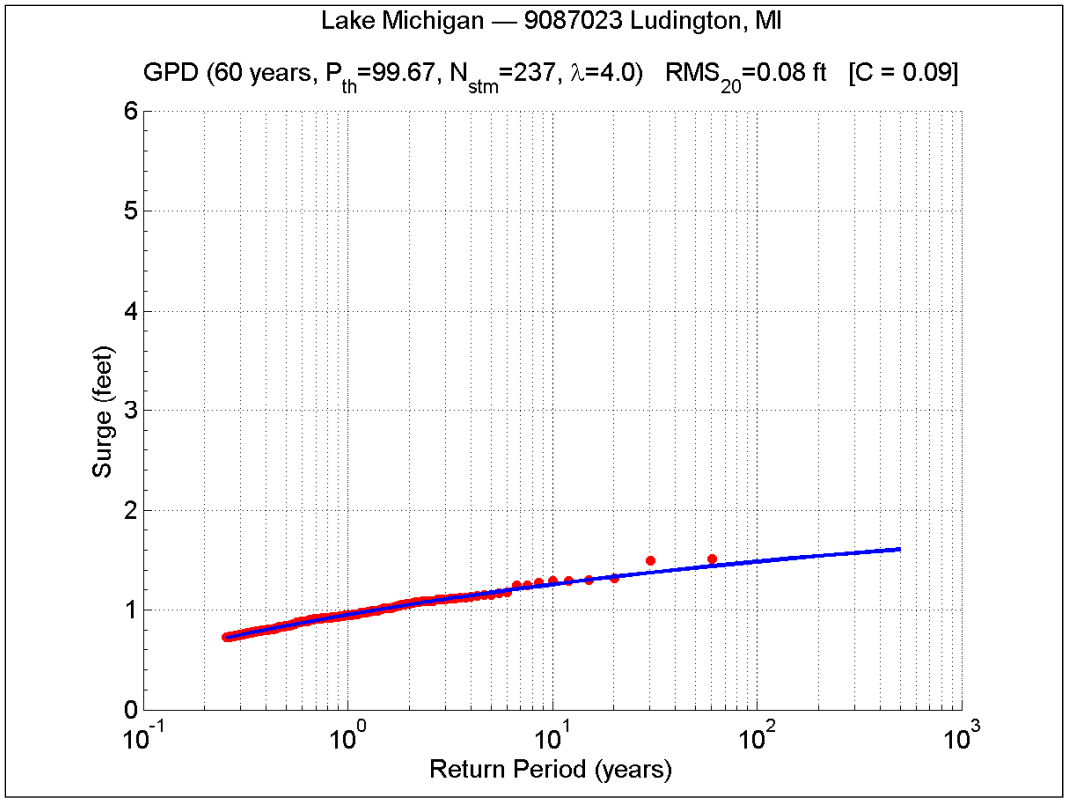



Figure 7. Return periods for surge from FSS for Ludington, MI.

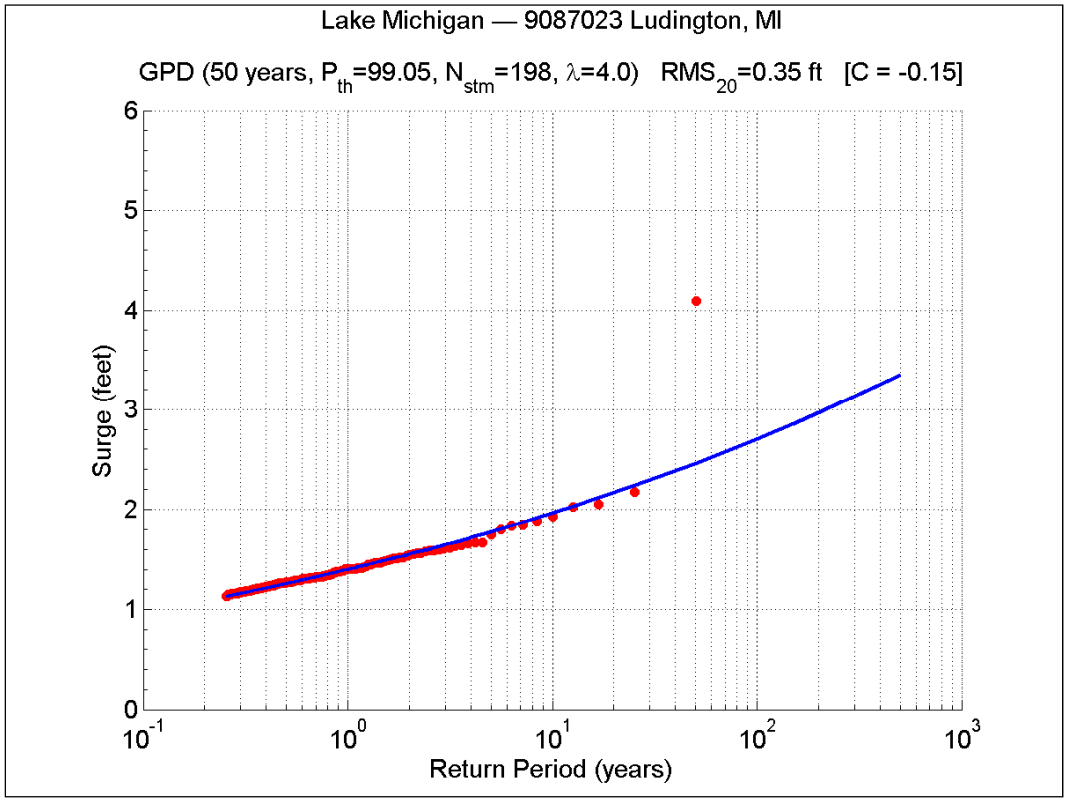



Figure 8. Return periods from surge estimates from wind surrogates for Ludington, MI.

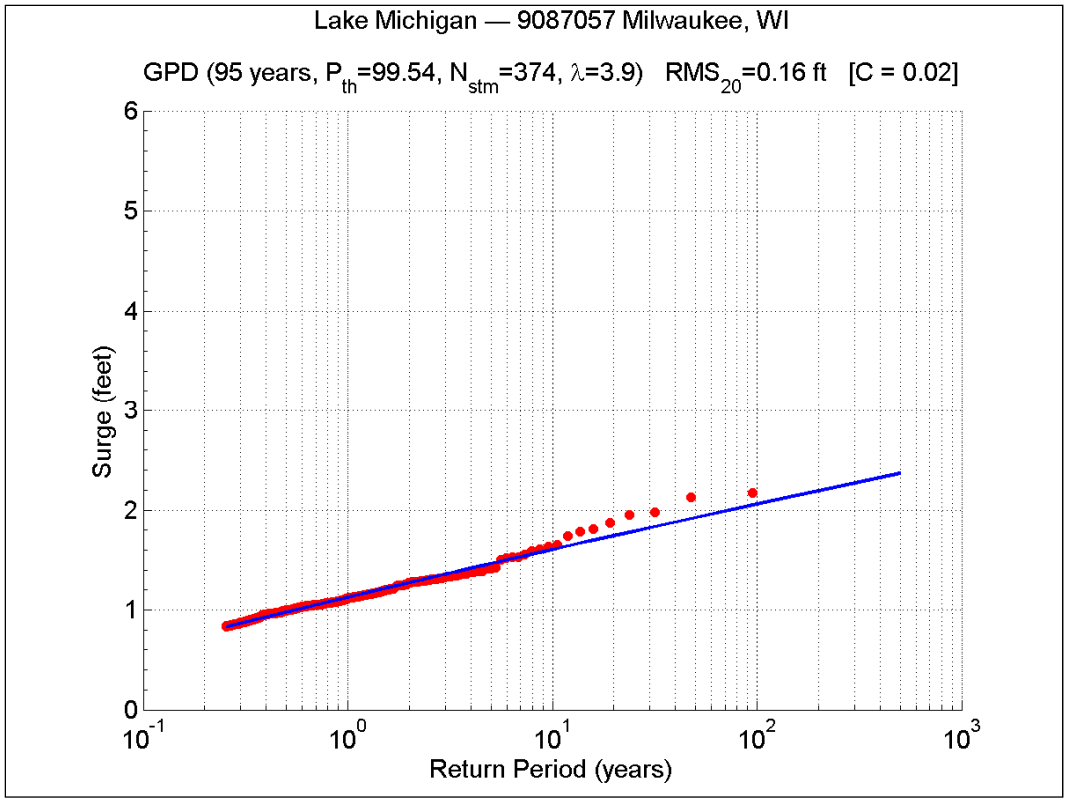



Figure 9. Return periods for surge from FSS for Milwaukee, WI.

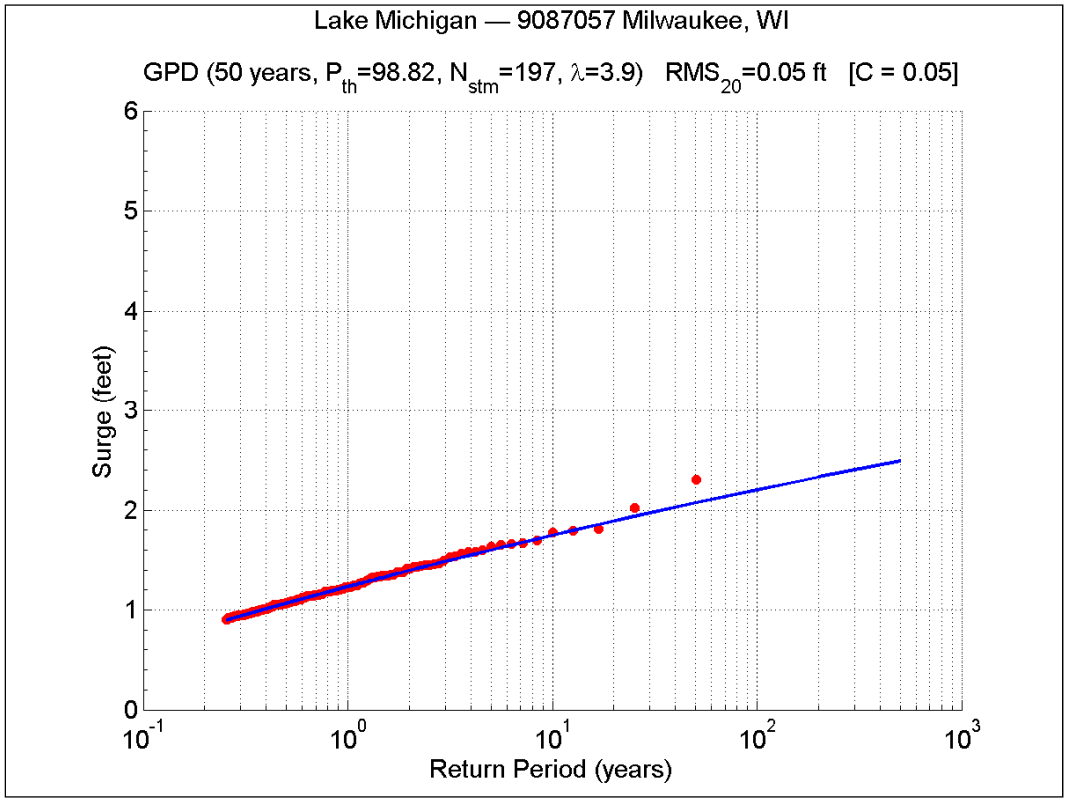



Figure 10. Return periods for surge estimates from wind surrogates for Milwaukee, WI.

Performing a POT analysis of the surrogate storm surge values resulted in completely different data sets due mainly to the overestimation of several surge events. As a result, it can be seen that the shapes of both the empirical distributions (red dots) and the GPD fits (blue curves) are clearly different in the extreme tails.

Additional goodness-of-fit parameters used in this study and cited in these figures and elsewhere in this report include:

1. Root-mean-square deviation (RMS), defined as:

$$RMS = \sqrt{\frac{\sum_{i=0}^{n} (X_i - Y_i)^2}{n}} \tag{11}$$

   RMS has the dimensions of the evaluated parameters (e.g., meters, feet). The RMS suffix included in the plots (e.g., RMS1, RMS10, RMS20) indicates the lowest return period for which RMS is computed. For example, RMS20 represents the root-mean-square deviation for all extreme events associated with a return period of 20 years or greater.

2. Concavity coefficient, *C*, is computed as the difference between the value of a log-linear function evaluated at $x_3$ and the value of the probabilistic distribution fit estimated at $x_3$. It can be written as:

$$C = \frac{f_L(x_3) - \hat{f}(x_3)}{|f_L(x_3)|} \tag{12}$$

   where

   $\hat{f}(x_3)$ = value of probabilistic distribution fit estimated at $x_3$ and

$$f_L(x_3) = m \log_{10}(x_3) + f_L(x_1) \tag{13}$$

$$m = \frac{f_L(x_2) - f_L(x_1)}{\log_{10}(x_2) - \log_{10}(x_1)} \tag{14}$$

   $f_L(x_1)$ and $f_L(x_2)$ =values of $f_L(x)$ evaluated at $x_2$ and $x_3$, respectively; $x_1 < x_2 < x_3$.

In terms of return periods (RP), the values of $x$ used in this study correspond to: $x_1 = RP(1\ year)$, $x_2 = RP(10\ years)$, and $x_3 = RP(500\ years)$.

## 5.3 Storm surge recommendations

The steady-state surrogate storm surge methodology outlined above was not accurate enough for the selection of storms. The dynamic effects involved in the estimation of storm surge are significant and not considering them is the main factor that led to these inaccuracies. Besides being steady-state, and ignoring the dynamic effects during storms, there are other factors that contributed to the inaccuracy, like the use of averaged water depths along fetches.

Therefore, in this study, measured water level data were utilized in the storm screening process to define FSS and CSS storm sets for storm surge. Alternately, if measured data are not available, a coarse-grid advanced circulation (ADCIRC) model could be used instead to reproduce surge time series and carry out the storm selection process.

# 6 Storm Surrogates – Waves

Currently, there are two types of wind-generated wave data readily available for Lake Michigan: NOAA National Data Buoy Center (NDBC) measurements, and Wave Information Study (WIS) hindcasts. Only two NDBC buoys are deployed in Lake Michigan. Moored buoy measurements, have several limitations including: short record lengths and the removal of buoys during winter season. Wave and over-water meteorological data available from NDBC are as follows: (1) buoy 45002: 9/1979 – present; (2) buoy 45007: 7/1981 – present. The WIS hindcasts are also limited by their record lengths. These data were only available from 1956 to 1997.

Due to limited wave data availability, an alternate surrogate method was sought to expand the wind-waves time series to cover the entire 1960-2010 period, thereby facilitating storm identification and selection. For this study, wind wave characteristics were computed based on methodologies used in ACES 2.0 software, and discussed in Leenkenecht et al. (1992), the CEM (USACE 2002), and in Smith (1991); Schwab and Morton (1984); Donelan (1980); Schwab (1978); Resio and Vincent (1977); and Cardone (1969).

The methodology that applies to the generation of waves within the confines of enclosed water bodies, such as lakes, is known as the restricted-fetch method. Wave heights and wave periods were estimated using assumptions of offshore and deepwater conditions. This approach made use of the land-based meteorological station wind data with 15-degree average bins. All of the necessary meteorological data corrections discussed previously were performed, including (1) overland to overlake adjustment, (2) equivalent neutral wind speed adjustment, and (3) wind speed averaging duration adjustment.

## 6.1 Wind-wave growth (restricted fetch)

The following are the equations utilized in the restricted-fetch wave computation process:

### 6.1.1 Primary wave development direction

This direction refers to a specific wind-fetch angle where the wind speed, and consequently wave height and wave period, is maximized as a function of fetch lengths at off-wind angles. This is accomplished by maximizing the fetch length, $F_\theta$, and the off-wind direction angle, $\theta$, as follows:

$$F_\theta^{0.28} = (\cos\theta)^{0.44} \tag{15}$$

Leenkenecht et al. (1992) recommend using this approach at 1-degree intervals. However, this was not feasible as part of this study since it required continuous wave estimates at 1-hour intervals throughout the 1960-2010 period. Therefore, the restricted-fetch approach was implemented using 15-degree directional bins. Further details are discussed by Smith (1991), based on concepts originally reported by Donelan (1980).

### 6.1.2 Wind-fetch vs. wind duration limitation

There are also limitations due to wind duration. If the wind duration, $t_{obs}$, is shorter than the duration required for fetch-limited wave conditions to develop then it would be considered duration-limited wave growth. This check is done by first computing the time required for fetch-limited conditions:

$$t_{fl} = 51.09 \frac{F^{0.72}}{g^{0.28} \hat{U}^{0.44}} \tag{16}$$

where

$t_{fl}$ = minimum duration required for fetch-limited conditions (sec)
$F$ = fetch length (m, ft)
$\hat{U}$ = fetch-parallel component of wind speed (m/sec, fps).

If $t_{obs} \geq t_{fl}$, then wave growth is considered fetch-limited. Conversely, if $t_{obs} < t_{fl}$, the duration-limited condition apply.

For fetch-limited conditions, wave height, $H_{fl}$, and wave period, $T_{fl}$, are computed as:

$$H_{fl} = 0.0015 \left( \frac{\hat{U}^2}{g} \right) \left( \frac{g F}{\hat{U}^2} \right)^{1/2} \tag{17}$$

$$T_{fl} = 0.3704 \left( \frac{\hat{U}}{g} \right) \left( \frac{g F}{\hat{U}^2} \right)^{0.28} \tag{18}$$

For duration-limited conditions, the following formulas are used:

$$H_{fl} = 0.000103 \left( \frac{\hat{U}^2}{g} \right) \left( \frac{g t_{obs}}{\hat{U}} \right)^{0.69} \tag{19}$$

$$T_{fl} = 0.082 \left( \frac{\hat{U}}{g} \right) \left( \frac{g t_{obs}}{\hat{U}} \right)^{0.39} \tag{20}$$

### 6.1.3 Fetch-limited vs. fully-developed condition

The final check performed in the fetch-limited methodology is to verify that both fetch-limited wave heights and periods do not exceed those of fully-developed conditions. First, fully-developed wave height, $H_{fd}$, and wave period, $T_{fd}$, need to be estimated from:

$$H_{fd} = 0.2433 \left( \frac{\hat{U}^2}{g} \right) \tag{21}$$

$$T_{fd} = 8.134 \left( \frac{\hat{U}}{g} \right) \tag{22}$$

The final wave height and wave period values are determined as the minimum of fetch-limited and fully-developed conditions:

$$H_{m0} = min\left( H_{fl}, H_{fd} \right) \tag{23}$$

where $H_{m0}$ = zero-moment significant wave height (m, ft), $T_p = min\left( T_{fl}, T_{fd} \right)$, and $T_p$ = spectral peak wave period.

Note that both non-convective (winter) and convective (summer) events are included as part of the surrogate wave results. However, it is questionable whether the surrogate wave methodology is actually able to adequately capture the wave effects for convective events, since they cannot be properly modeled with currently available wind forcing data due to insufficient temporal and spatial resolution.

## 6.2 Wind-wave results

Table 4 shows statistics comparing the surrogate wave estimates to the WIS hindcast waves for each of the NOAA gages listed in Table 2. These statistics are: (1) average difference between surrogate waves and hindcast waves; (2) average percentage difference (differences divided by hindcast wave values); (3) RMS; and (4) normalized RMS, or NRMS (RMS divided by the average hindcast wave value).

Figures 11 and 12 show the comparison between hindcast and estimated wave surrogate values for the top 20 and top 100 storms for the Ludington gage, respectively. It can be seen that the spread in these plots is less and the correlation is higher compared to the surrogate storm surge results (Figures 3 and 4). Adequate correlation was found between extreme WIS hindcast waves and the surrogate waves.

Table 4. Statistics comparing surrogate wave results to WIS hindcasts.

| Top 20 wave events at NOAA water level gage | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Station No. | 9087023 | 9087031 | 9087044 | 9087057 | 9087068 | 9087072 | 9087096 | Average |
| Diff (ft) | -0.31 | -0.76 | 0.08 | 0.35 | -0.05 | 0.07 | 0.12 | -0.07 |
| Diff (%) | -1.6 | -4.1 | 1.3 | 2.2 | 0.0 | 0.9 | 1.5 | 0.0 |
| RMS (ft) | 1.84 | 2.53 | 1.33 | 1.98 | 1.55 | 1.67 | 1.80 | 1.81 |
| NRMS (%) | 10.8 | 13.8 | 7.8 | 12.1 | 10.7 | 12.0 | 10.8 | 11.2 |
| Top 100 wave events at each NOAA water level gage | | | | | | | | |
| Station No. | 9087023 | 9087031 | 9087044 | 9087057 | 9087068 | 9087072 | 9087096 | Average |
| Diff (ft) | 1.16 | -1.26 | 1.09 | 0.68 | 1.25 | 0.49 | 0.95 | 0.62 |
| Diff (%) | 9.6 | -7.7 | 9.4 | 5.4 | 11.5 | 4.8 | 8.1 | 5.9 |
| RMS (ft) | 2.48 | 2.53 | 2.18 | 2.72 | 2.08 | 2.39 | 2.45 | 2.41 |
| NRMS (%) | 17.8 | 16.2 | 16.4 | 20.5 | 17.6 | 21.2 | 18.7 | 18.3 |

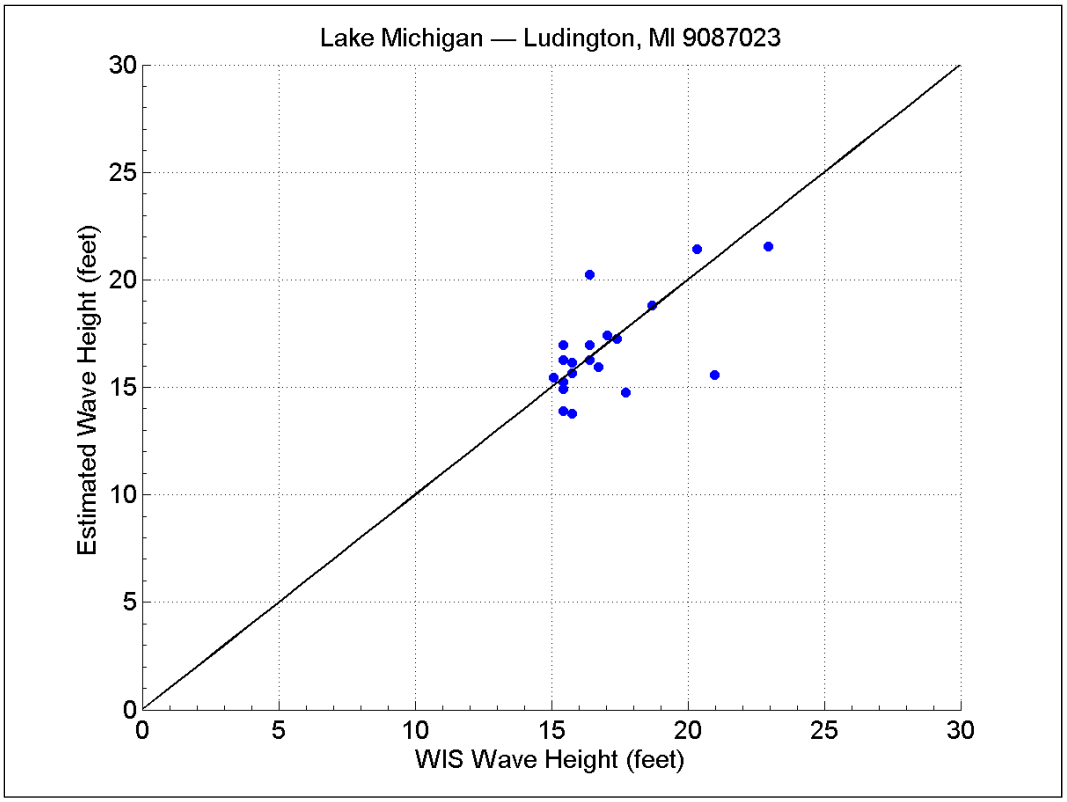



Figure 11. Top 20 wave events for Ludington, MI gage.

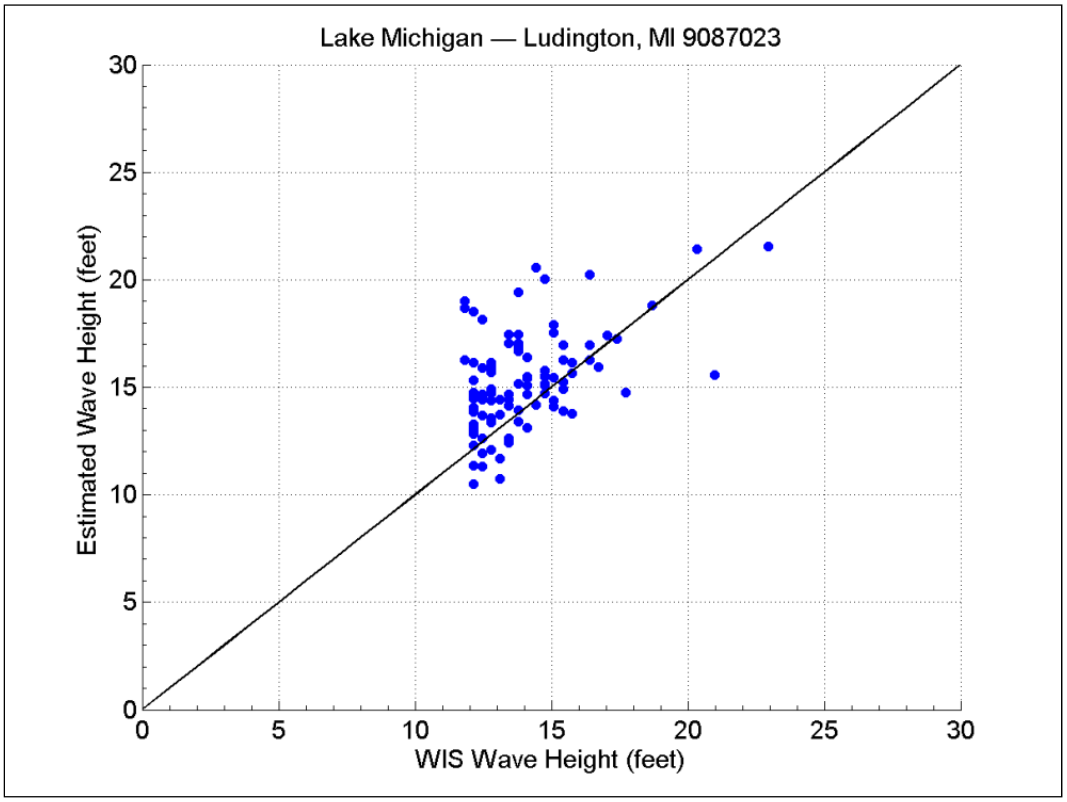



Figure 12. Top 100 wave events for Ludington, MI gage.

Figure 13 shows the return period plot of wind wave surrogate estimates for the Ludington gage. The return period plot for hindcast data is not shown because the years in record do not coincide. WIS hindcast data is limited to the 1956-1997 period, while the wind data used to generate surrogate waves span the 1960-2010 period. The disparity in record lengths and years of availability lead to different GPD fits, regardless of how accurate ACE/CEM-method estimates are.

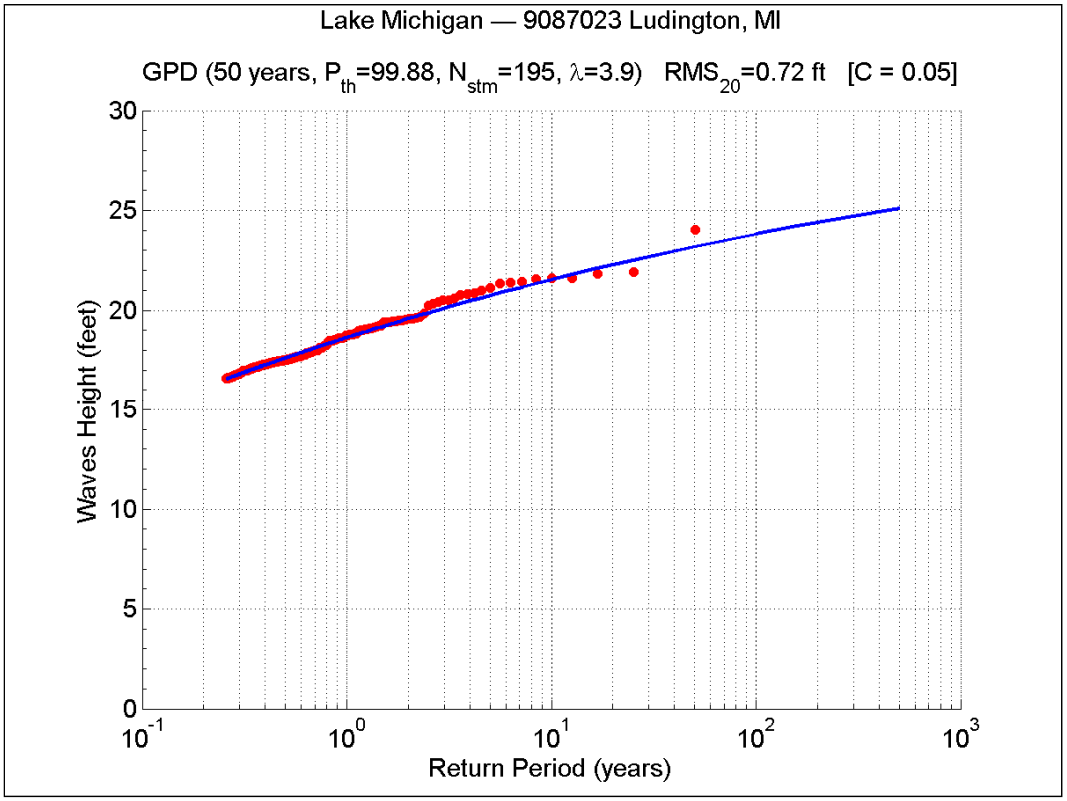



Figure 13. Return periods for wave estimates from wind surrogates for Ludington, MI.

Figures 14 and 15 show the comparison between hindcast and estimated surrogate wave values for the top 20 and top 100 storms for the Milwaukee gage, respectively. Similar to the Ludington location, the correlation between extremal hindcast waves and ACE/CEM-method estimated waves is reasonable to justify use of surrogate waves in the screening process.

Figure 16 shows the return period plot of wind wave estimates for the Milwaukee gage. As noted for Ludington, the return period plot for WIS hindcast data is not shown due to discrepancies in record lengths (1960-2010 for surrogate waves vs. 1956-1997 for WIS hindcasts).

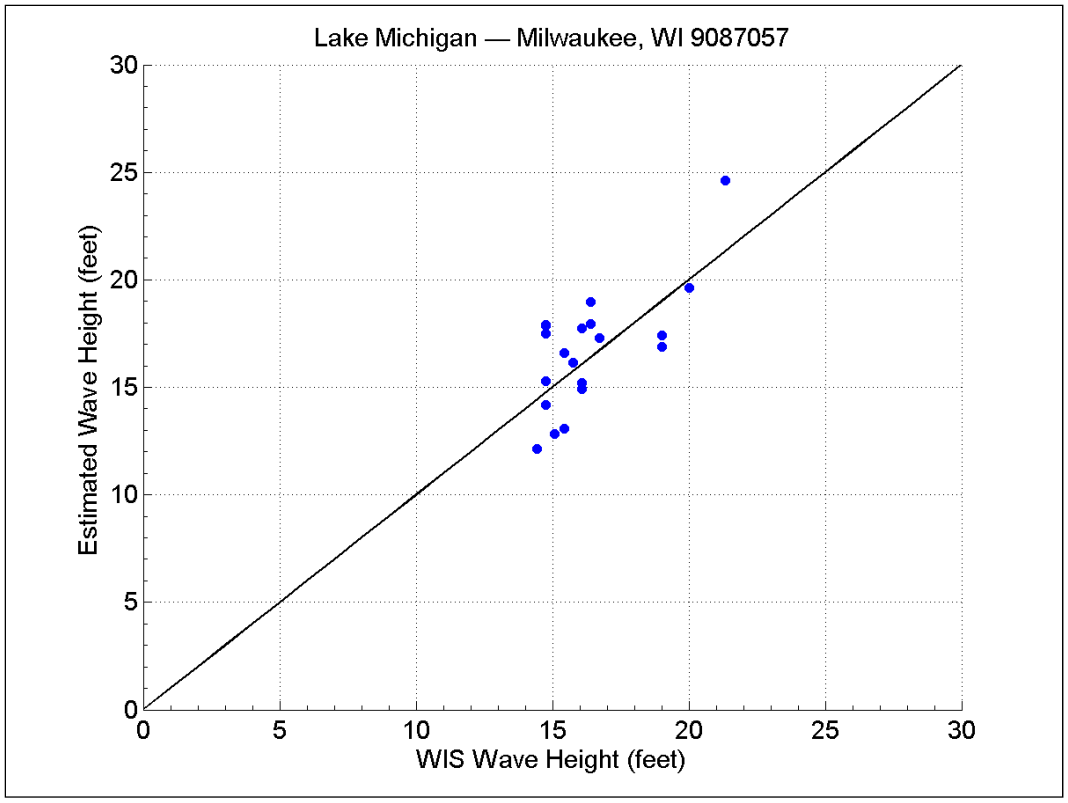



Figure 14. Top 20 wave events for Milwaukee, WI WL Gage.

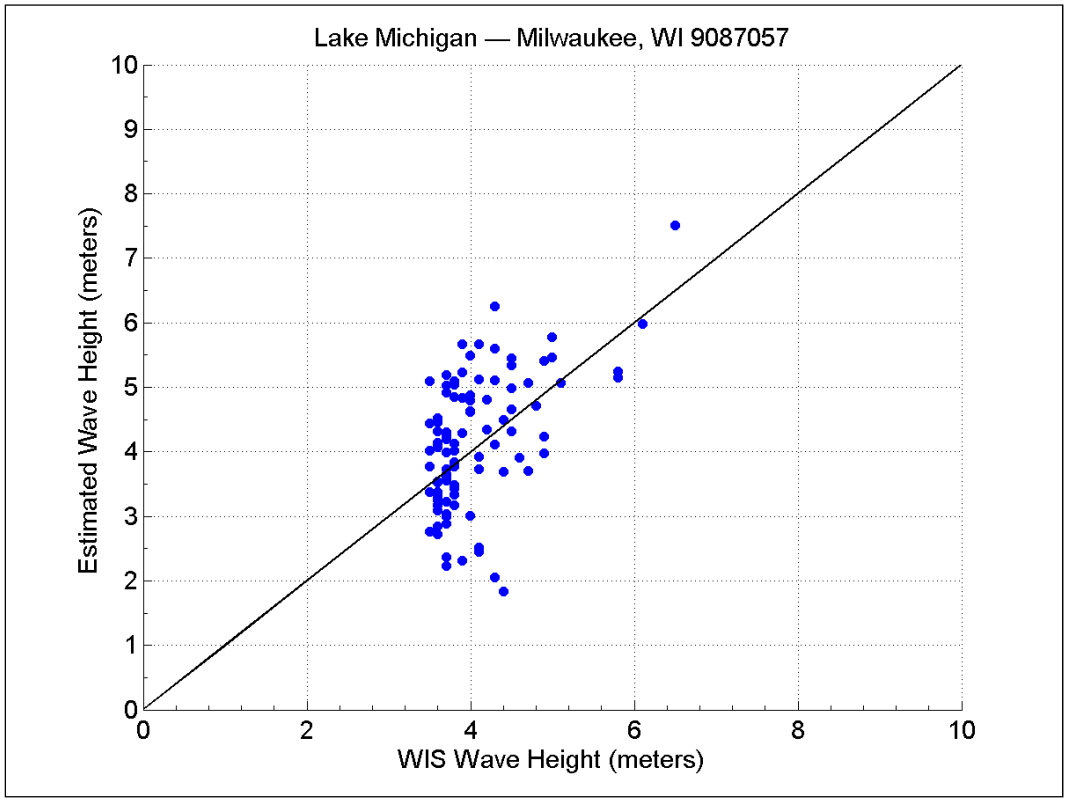



Figure 15. Top 100 wave events for Milwaukee, WI WL Gage.

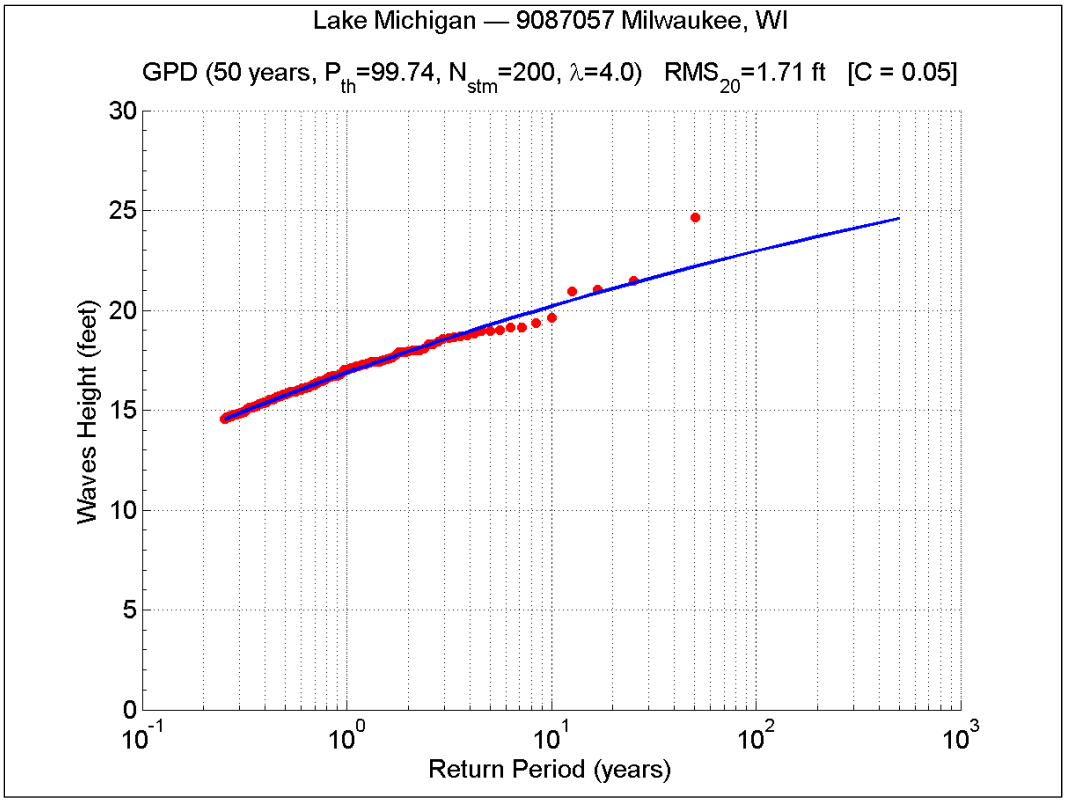



Figure 16. Wave estimates from wind surrogates for Milwaukee, WI.

## 6.3 Wind-wave recommendations

Comparisons of surrogate wave heights to WIS hindcasts for the top 20 and top 100 wave events showed that, on average, NRMS deviations are roughly around 10 percent and 20 percent, respectively. Likewise, on average, the percentage differences between surrogate waves and hindcasts are: 0 percent for the top 20, and six percent for the top 100 events. The use of ACES/CEM wind-wave generation methodology is an adequate alternative, in the absence of measured or hindcast data, to estimate waves as part of the storm selection and screening process.

# 7 Storm Sample Size

The adequacy of the statistical approach and the required sample size for a 50-year record was evaluated. The evaluation of storm selection and sample size requires comparing two probability distributions: (1) a distribution derived from the full storm set (FSS), and (2) a distribution derived from a sample storm set. The sample storm set is considered contained entirely within the FSS. Herein, the sample storm set is termed the composite storm set (CSS) because it is a composite of samples from all of the gage locations around the lake. The FSS distribution represents the true water level distribution at a particular location, or at least as close as one can get to the true distribution, based on the period of record used.

The main objective of this task is to determine the minimum number of storms that should be sampled from each gage location, and included in the CSS, to adequately represent the true water level distribution at all locations throughout the lake. This is done here by comparing the storm still water level (SWL = lake level + surge) exceedance distributions derived from the CSS to the distributions corresponding to the FSS, on a gage-by-gage basis.

The analysis, performed using SWL only, with no consideration of wave influence, is summarized as follows:

1. CSS are created by selecting unique storm events from different locations. For this analysis the CSS were generated by sampling the top $N = 1, \ldots, 100$ surge events (difference between SWL and lake level) from each of the locations corresponding to the nine NOAA water level gages; in the case of Lake Michigan, this initial sampling procedure results in 100 different CSS, ranging from 9 to 900 storms in each set.
2. Identify and replace duplicate storms. When duplicate storms are identified, meaning storms that are affecting more than one gage, the storm with the highest ranking (location-wise) is retained; the remaining duplicate storm or storms are removed from the CSS and replaced with the next storm in rank from the same gage location.
3. Return period plots were generated for the SWL corresponding to the sampled storms. The results derived from the CSS were compared to those from the FSS. This was done by evaluating how well the one percent and

0.2 percent annual chance values as estimated from SWL-CSS matched those corresponding to SWL-FSS. The FSS at each gage location is defined by all events that exceed the adopted storm surge threshold at that location. Each threshold value was iteratively adjusted (on a location-by-location basis) until $\lambda$= 4-5 was obtained; that is, the threshold value that is adopted for each location should result in the sampling of four to five storm events per year, on average.

The minimum number of storms that should be sampled per gage to match the extreme tail of the distributions and achieve accurate one percent and 0.2 percent annual chance SWL values varies slightly from location to location. Figure 17 shows the differential SWL (i.e., difference between SWL-FSS and SWL-CSS values) for Ludington. The blue curve represents the differences in one percent annual chance SWL, while the red curve represents the differences in 0.2 percent annual chance SWL. Likewise, Figure 18 shows the differential SWL for Milwaukee. In general, the SWL values corresponding to these low-frequency events are stable when 9- 22 storms are sampled per gage. Beyond these numbers additional variability is observed, roughly 20 percent or more, because this leads to the sampling and inclusion of non-extreme events which tend to change the shapes of the GPD fits. Therefore, our recommendation is to limit the number of sampled storm events per gage within this 9-22 range.

All gage locations considered, it was found that the minimum number of storms that must be sampled to reasonably match the one percent and 0.2 percent annual chance SWL values is around 10 per gage (roughly equivalent to a 100-storm set). However, this analysis only evaluated differences in the upper tail of the SWL distributions (low-frequency events). Additional analyses discussed in the upcoming section showed that a larger number of storms must be sampled per gage to match both the low- and high-frequency portions of the distribution and to obtain the correct GPD shapes.

## 7.1 GPD fits vs. storm sample size

Using the 100-storm CSS to develop the GPD-based return period plots yielded satisfactory results at the upper tail of the distributions; however, the overall shapes of the distributions were unsatisfactory. While the 100-storm set adequately captures the one percent and 0.2 percent annual chance SWL values, the distribution shapes resulted in most high-frequency events being underpredicted.

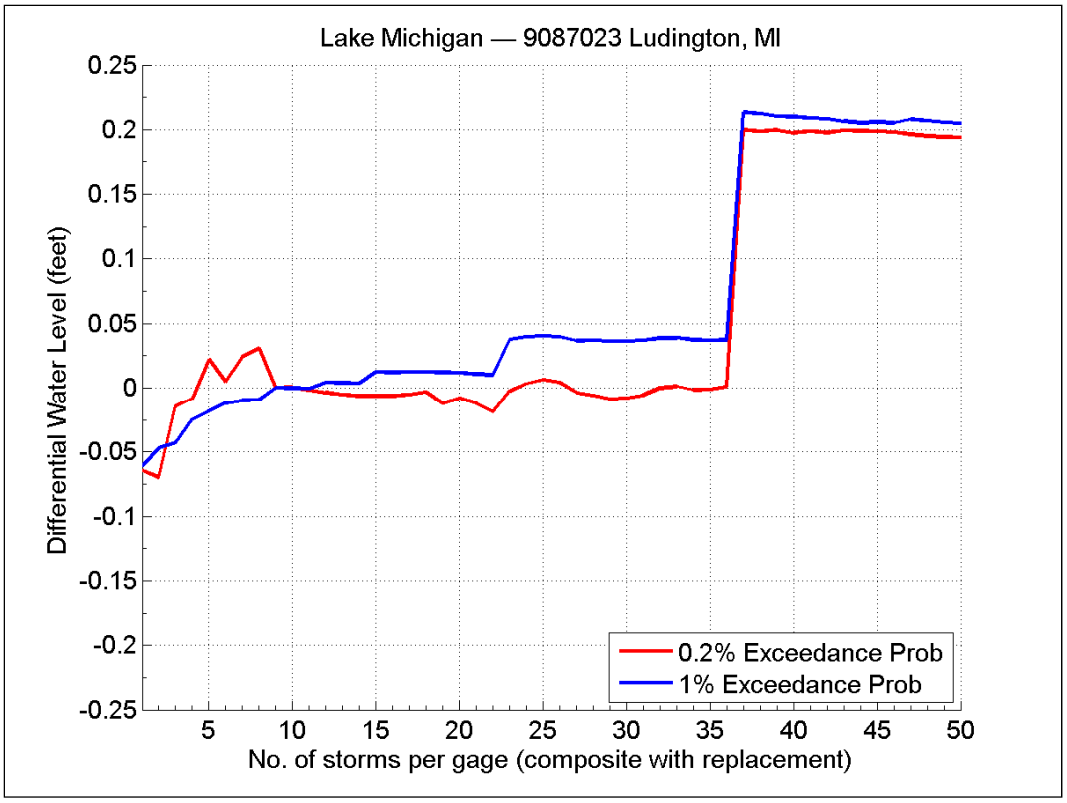



Figure 17. Difference between FSS and CSS vs. number of storms for Ludington, MI.

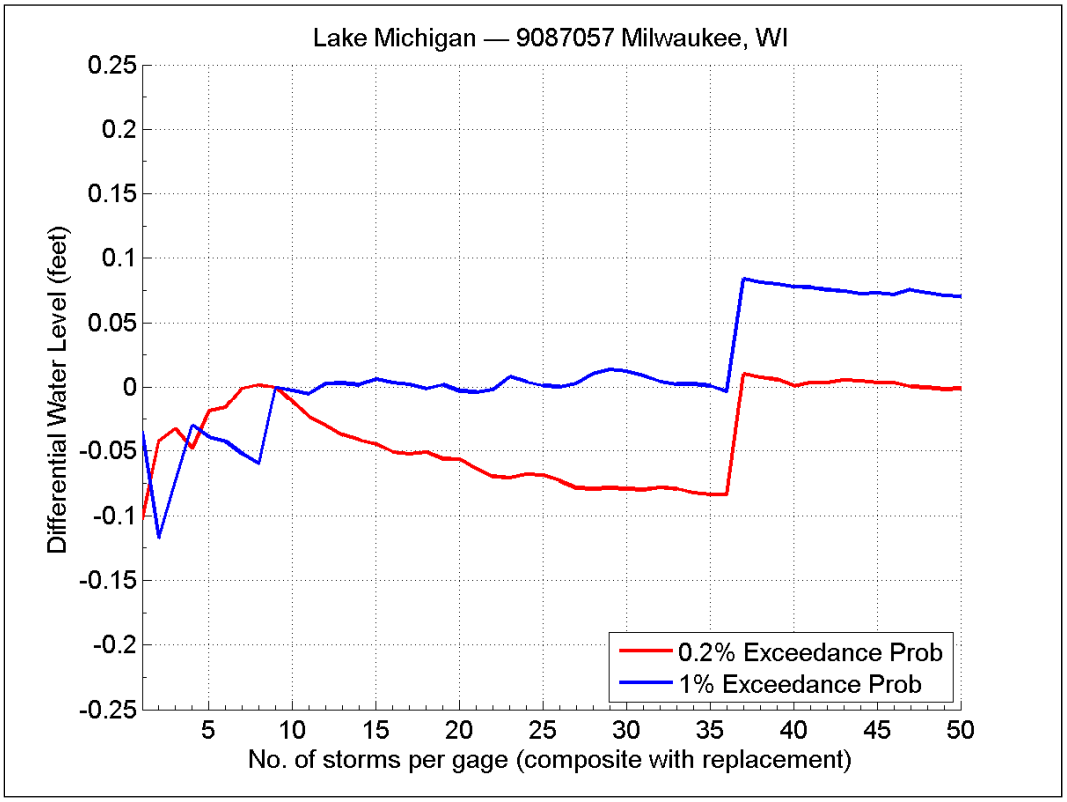



Figure 18. Difference between FSS and CSS vs. number of storms for Milwaukee, WI.

To correct the problem associated with high-frequency events, additional in-depth assessments of SWL distributions were performed. These analyses consisted of evaluating SWL values of all frequencies as well as overall GPD shapes. Figures 19 through 21 show the SWL Return Period (RP) plots corresponding to CSS comprised of 100, 150 and 175 storms, respectively, for the Ludington gage. Likewise, Figures 22-24 show the same plots for the Milwaukee gage. Differences in the high-frequency tail of the 100-storm and 150-storm distributions at each location are clearly evident.

The results showed that the distribution shapes generated with 150 storms are much closer to the FSS distribution than those distributions generated with just 100 storms. Thus, increasing the sample size from 100 to 150 total storms significantly improves the shape of the distributions and provides adequate fits to high frequency events (e.g., 2-year return period). Differences between actual and estimated SWL for these events were reduced to around 0.20 to 0.40 feet, depending on the station. Therefore, in the case of Lake Michigan, the ideal minimum number of total storms that should be sampled to generate the CSS is approximately 150 storms, or roughly 16-17 storms per gage. Further increasing the total sample size beyond 150 storms provided minimal benefits in terms of SWL distribution accuracy.

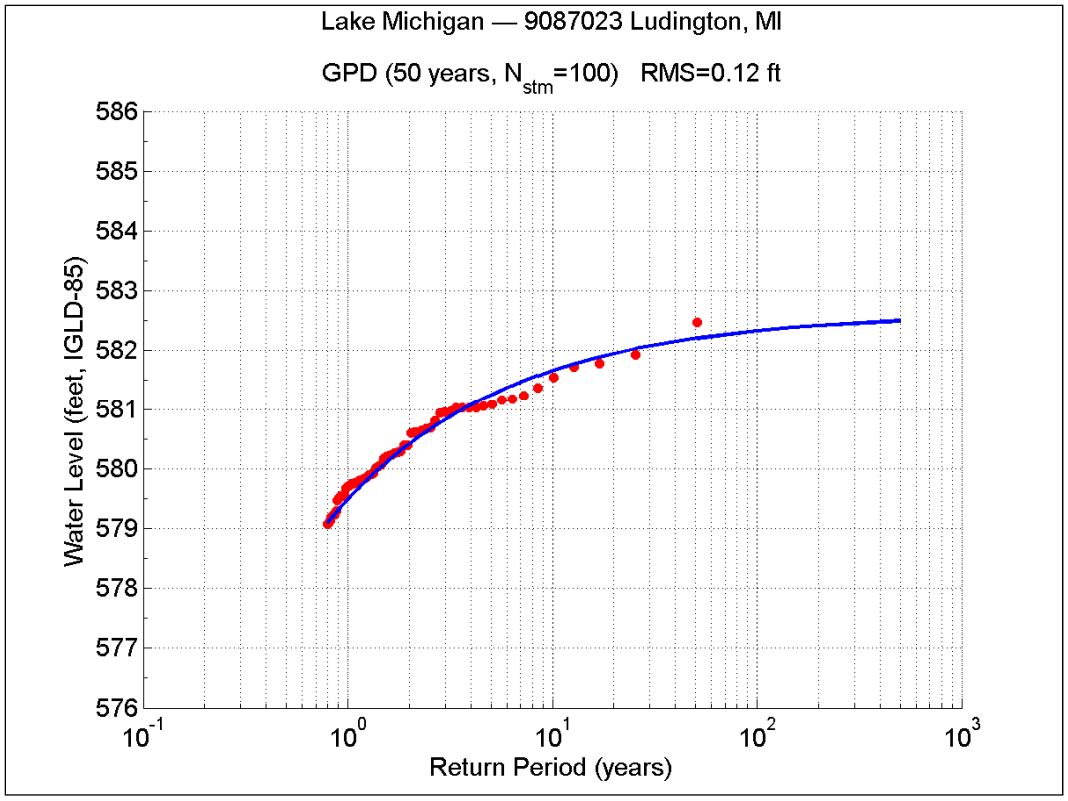



Figure 19. SWL RP plot from 100-storm CSS for Ludington, MI gage.

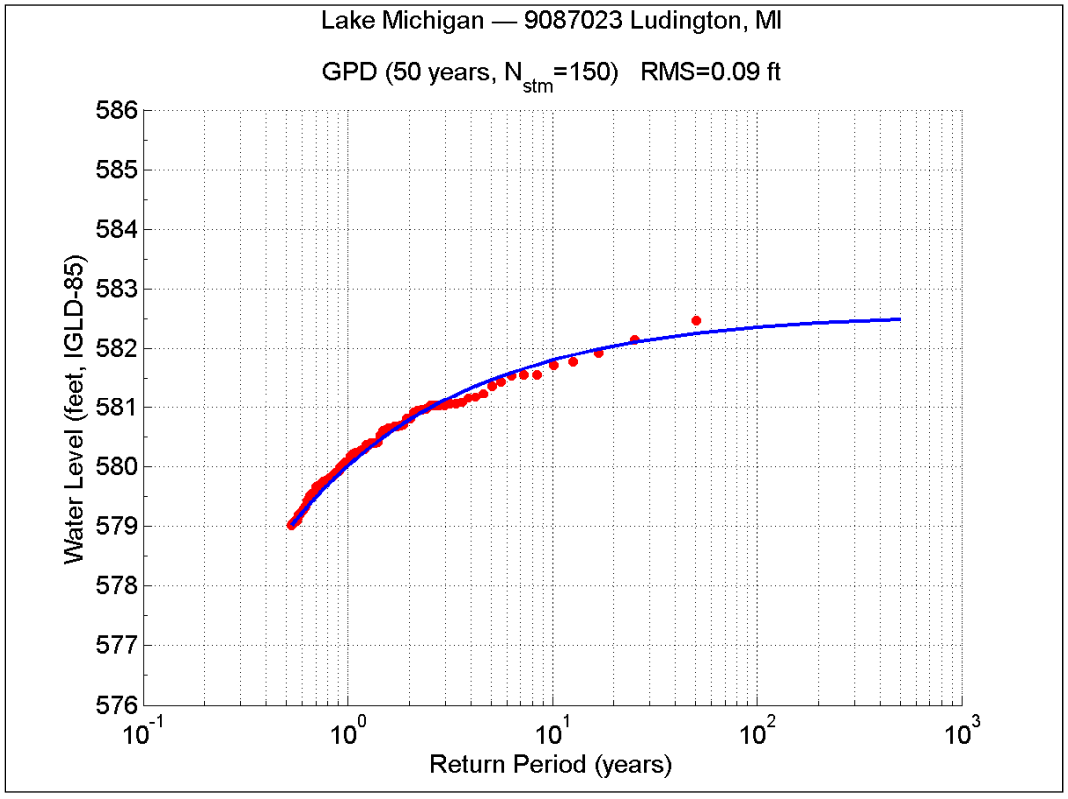



Figure 20. SWL RP plot from 150-storm CSS for Ludington, MI gage.

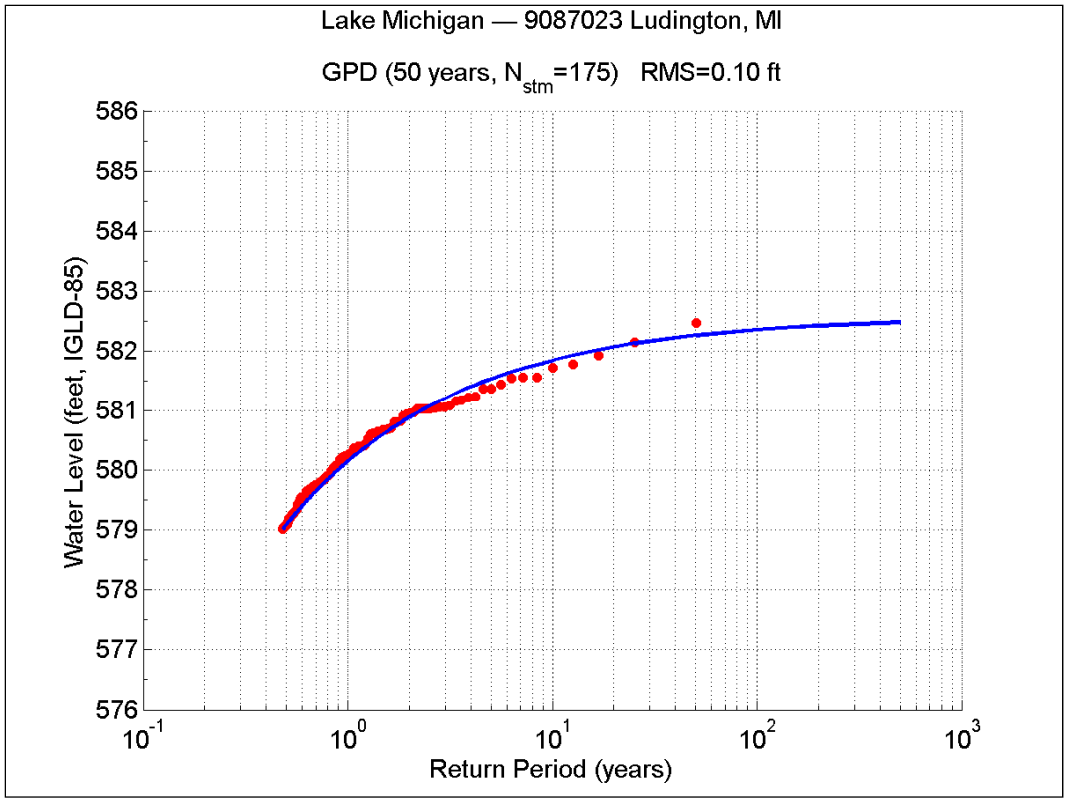



Figure 21. SWL RP plot from 175-storm CSS for Ludington, MI gage.

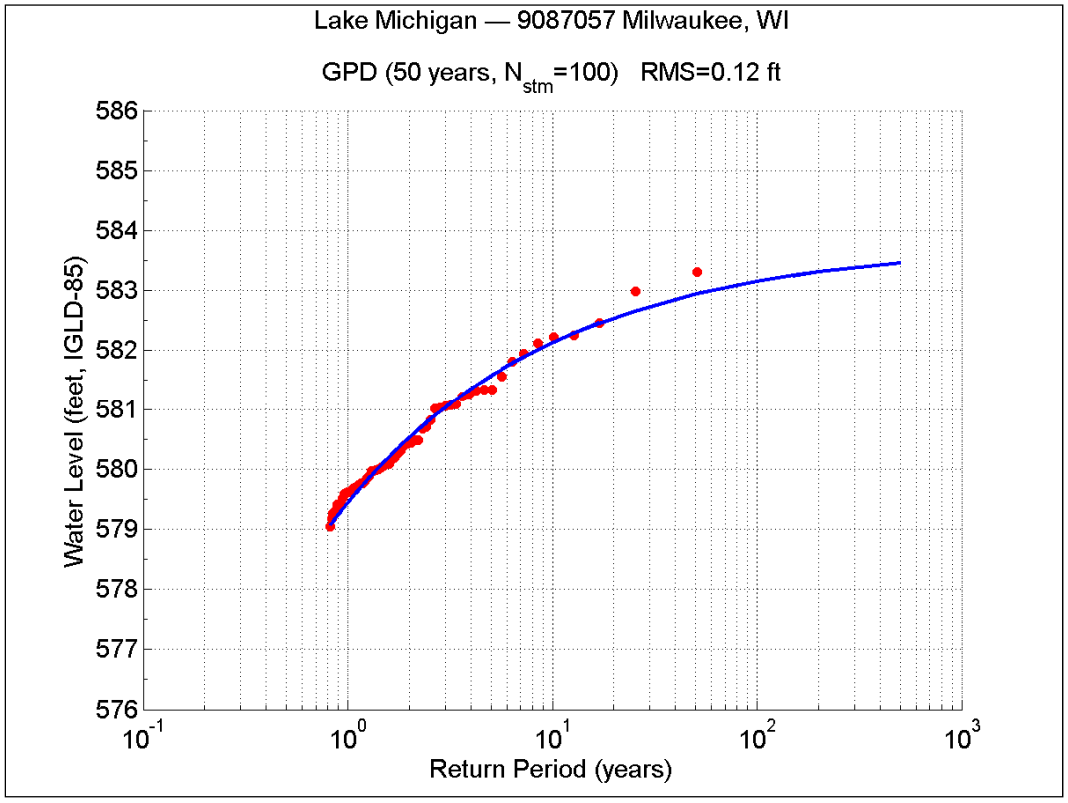



Figure 22. SWL RP plot from 100-storm CSS for Milwaukee, WI gage.

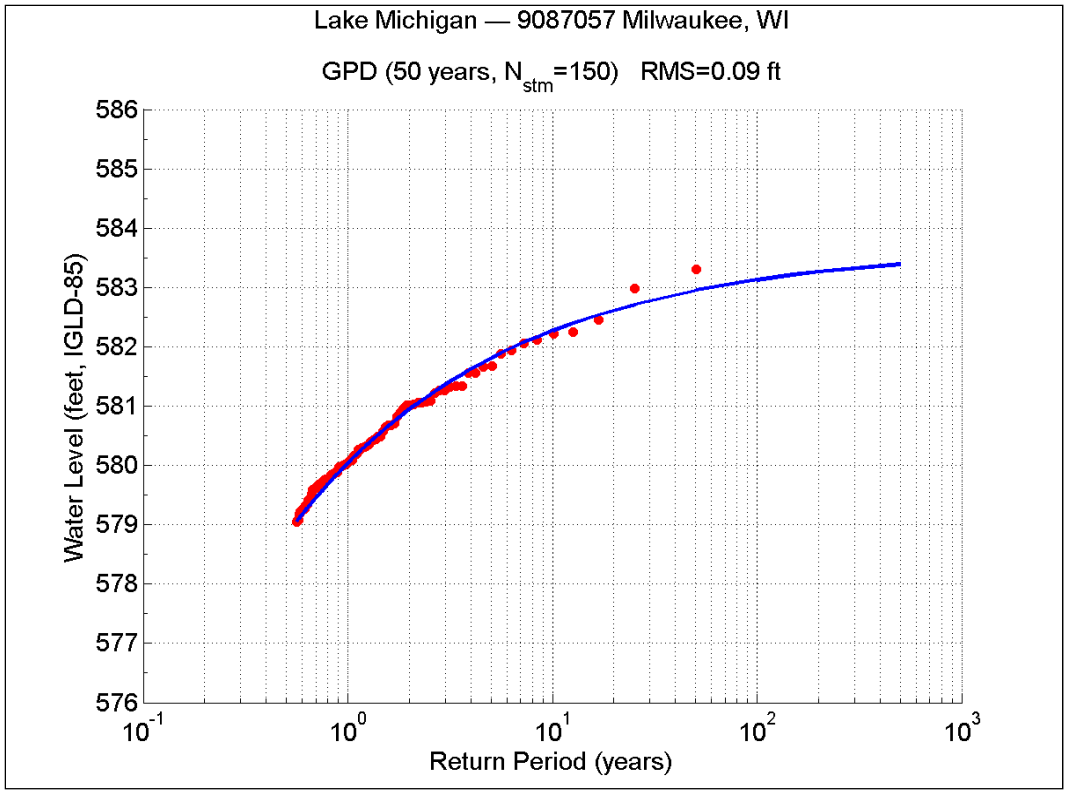



Figure 23. SWL RP plot from 150-storm CSS for Milwaukee, WI gage.

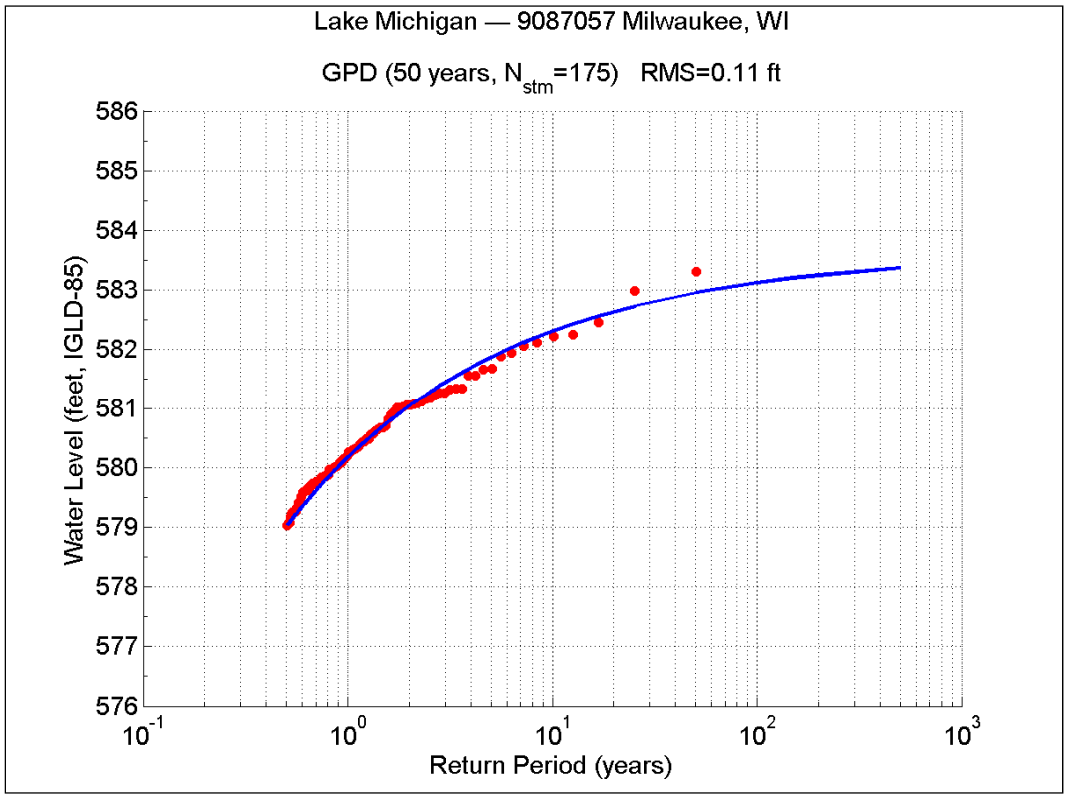



Figure 24. SWL RP plot from 175-storm CSS for Milwaukee, WI gage.

# 8 Sampling Across High and Low Lake Levels

A critical issue in developing a storm sampling approach is whether the events that end up being sampled are representative of the entire record length. In the case of the Great Lakes, and of Lake Michigan, the main question is if the sampling is actually being done across both high and low lake-wide levels. If the storm sampling is done solely on high lake levels, for example, this would result in ill-shaped exceedance distributions and could bias the one percent and 0.2 percent annual chance water levels. To avoid introducing bias into BFEs probabilities, it is important for storm sampling to be done to properly reflect the distribution of lake levels.

This issue of sampling across all lake levels was further assessed through a resampling analysis. Different resampling methods were utilized, such as nearest neighbor resampling (NNR), cubic spline and linear interpolation. These methods are schemes that can be used to emulate signals (frequency domain) and time series (time domain); in this case the time series to be emulated is the long-term lake level. The intent is to sample enough storm events across all lake levels in order for the resampled time series to be an accurate representation of the original time series.

The observations used in this analysis were the lake-wide (average of all gages) maximum hourly level for each month that a sample storm was identified. These monthly maximum levels were used because they account for, both, lake-levels and storm surge elevations. Statistics were computed to determine the minimum number of storms that needed to be sampled per gage to assure adequate sampling across all lake levels; computed statistics included variance, standard deviation, and model error tests using RMS deviation. The statistics were computed from all the monthly maximum lake-wide level values for the entire 50-year period between 1960-2010. Therefore, the 600 lake-wide monthly values (50 years x 12 months) of the original time series were compared to the resampled time series.

It was determined that the NNR method provided the optimal results since this method was able to closer emulate the SWL time series. The cubic spline resampling method yielded less than desired results because

spurious peaks biased the lake level statistics. The linear interpolation method did not produce statistics similar to the lake levels.

Using the NNR method, the gap between two consecutive observations ($O1$, $O2$) is filled by first dividing the gap in two segments of equal length; the value of $O1$ is assigned to the first segment, while $O2$ is assigned to the second segment. This process is repeated to fill gaps between all observations. The final result is an aliased resampled time series of the original lake-wide level data. No anti-aliasing technique was used as part of this analysis.

Figure 25 shows the percentage difference between the variance of the lake-wide SWL time series (for the 1960-2010 period) and the variance of the NNR-generated time series, as a function of the number of storms sampled per gage. Likewise, Figure 26 presents the same analysis but employing standard deviation instead of variance. In general, both differential variance and differential standard deviation are minimized when 15 to 22 storms are sampled per gage. The values shown on both of these figures (i.e., variance and standard deviation) are averages computed from all nine water level gages.

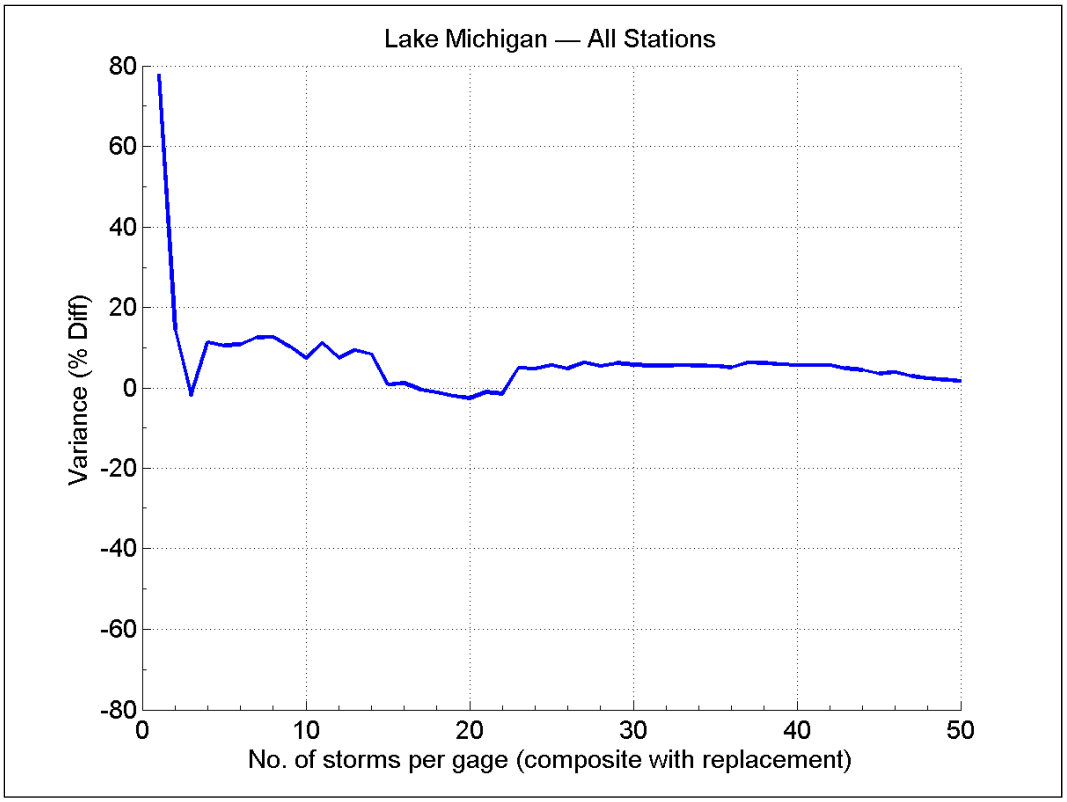



Figure 25. Differential SWL variance vs. number of storms sampled per gage.

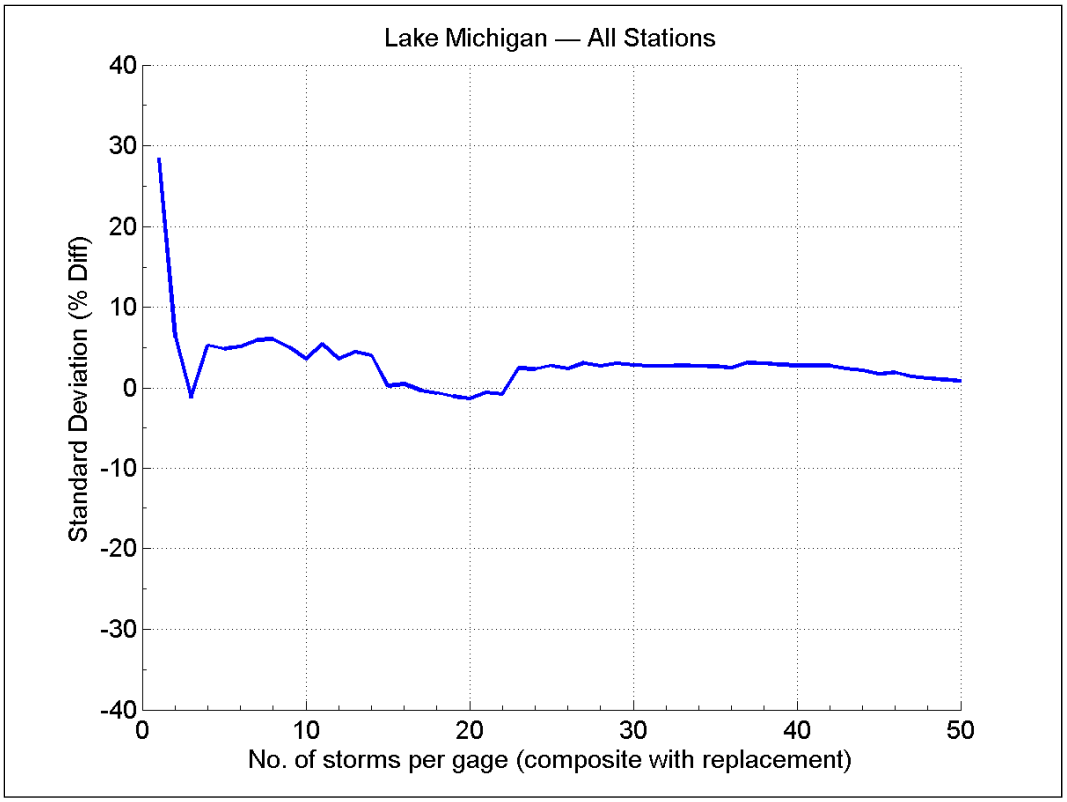



Figure 26. Differential lake level standard deviation vs. number of storms sampled per gage.

Therefore, from this resampling analysis it can be concluded that at least 15 storms per gage (or 135 total storms) are necessary for adequate emulation of the long-term water level signal and its distribution. This is in agreement with the storm sample size analysis discussed in Chapter 7 (Storm Sample Size) which required a CSS of roughly 150 storms.

Based on the model error analysis (see Figure 27), the RMS deviation (between the lake-wide water level time series and the NNR-resampled time series) corresponding to the 150-storm set is approximately 0.40 feet. Increasing the total number of storms to 175 barely decreased the RSM deviation by only 0.05 feet. This further justifies the 150-storm sample size.

The distribution of storms contained in the 100-storm CSS across all lake levels for the 1960-2010 period are shown for the Ludington and Milwaukee gages in Figures 28 and 29, respectively. Similarly, distributions of storms for the 150-storm CSS are shown in Figures 30 and 31. Also illustrated in these plots are the 1960-2010 monthly maxima water levels (red lines) and the NNR fits (blue lines).

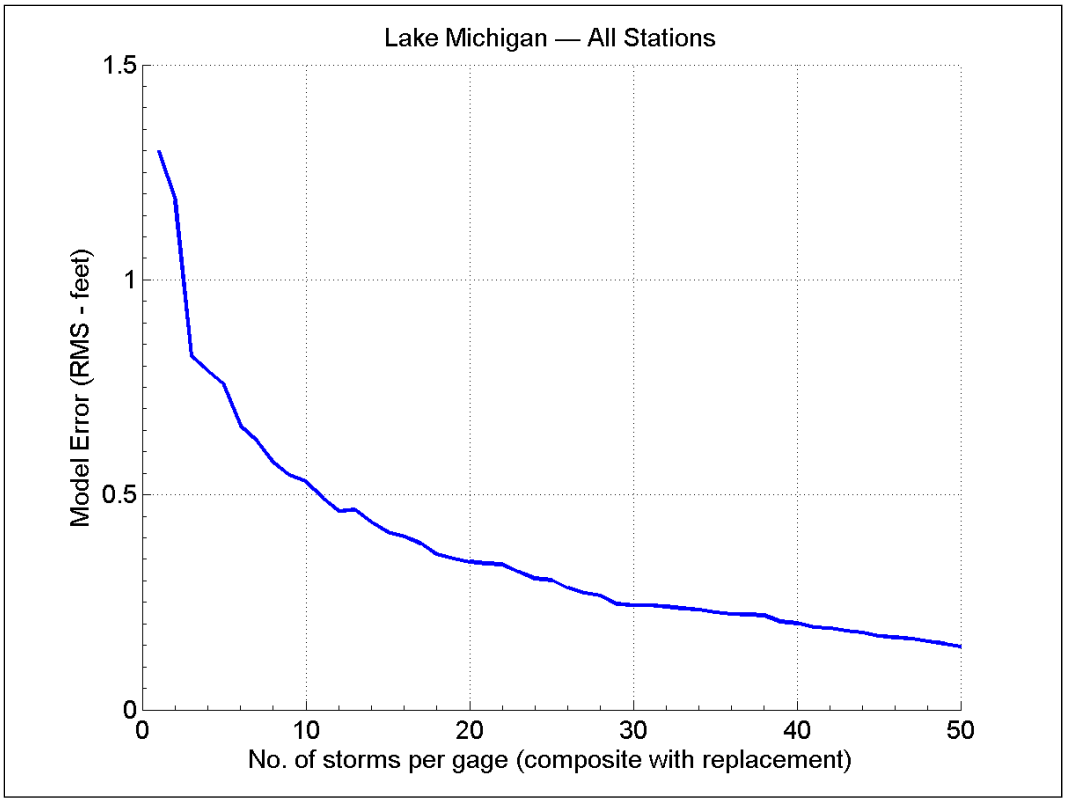



Figure 27. Lake level model error (RMS) vs. number of storms sampled per gage.

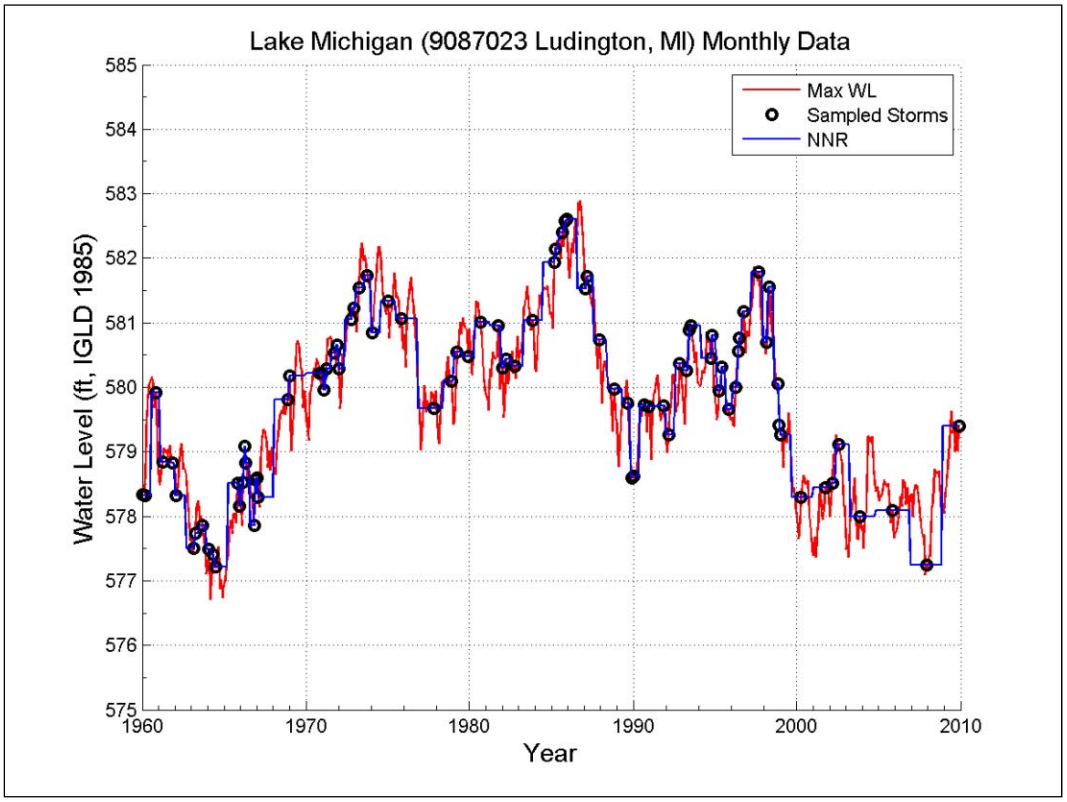



Figure 28. Distribution of 100-storm CSS across lake levels for Ludington, MI.

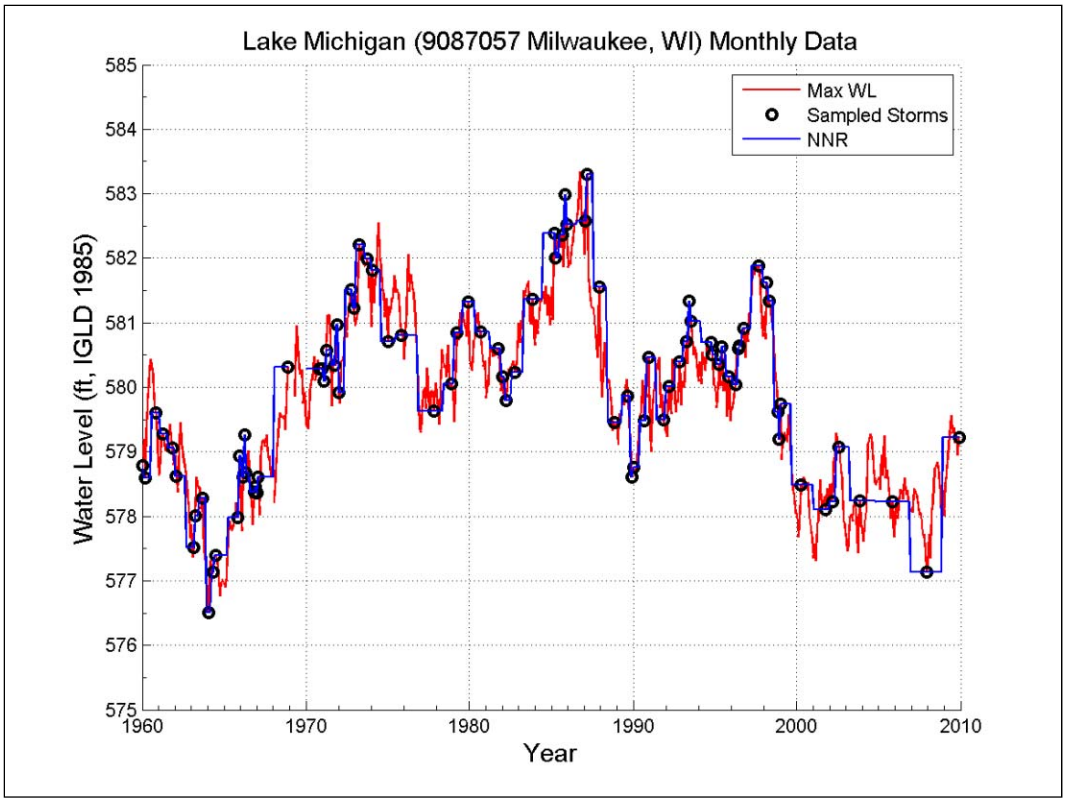



Figure 29. Distribution of 100-storm CSS across lake levels for Milwaukee, WI.

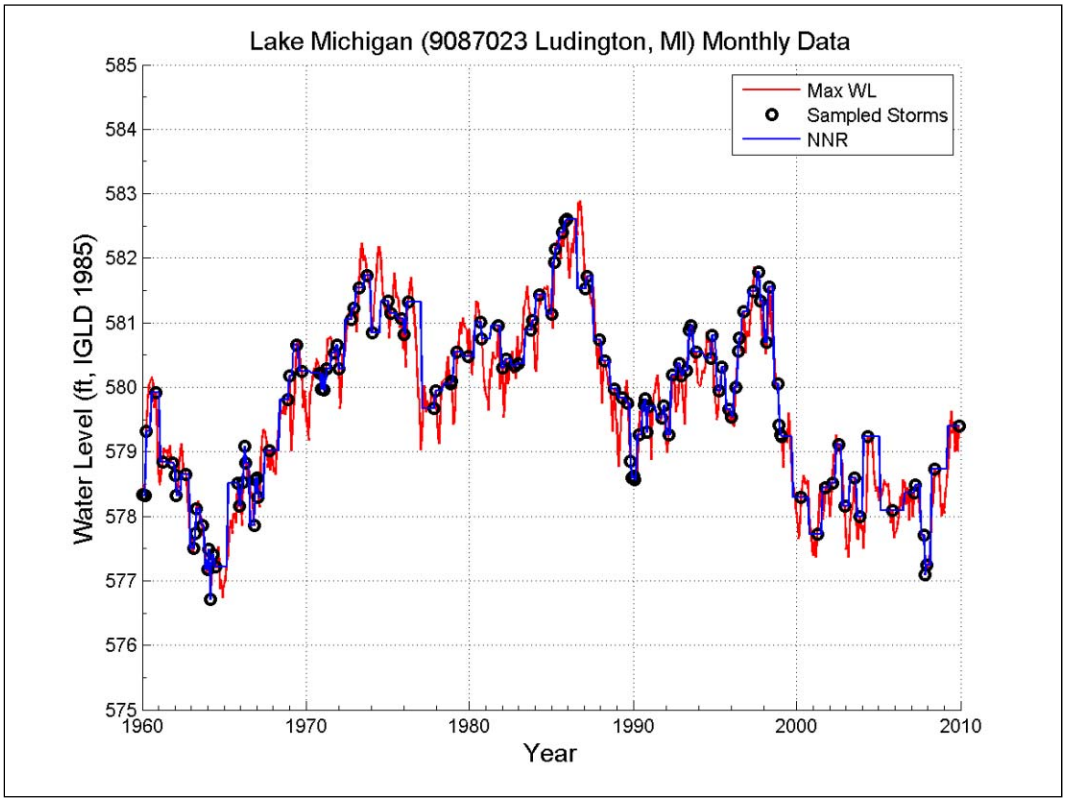



Figure 30. Distribution of 150-storm CSS across lake levels for Ludington, MI.

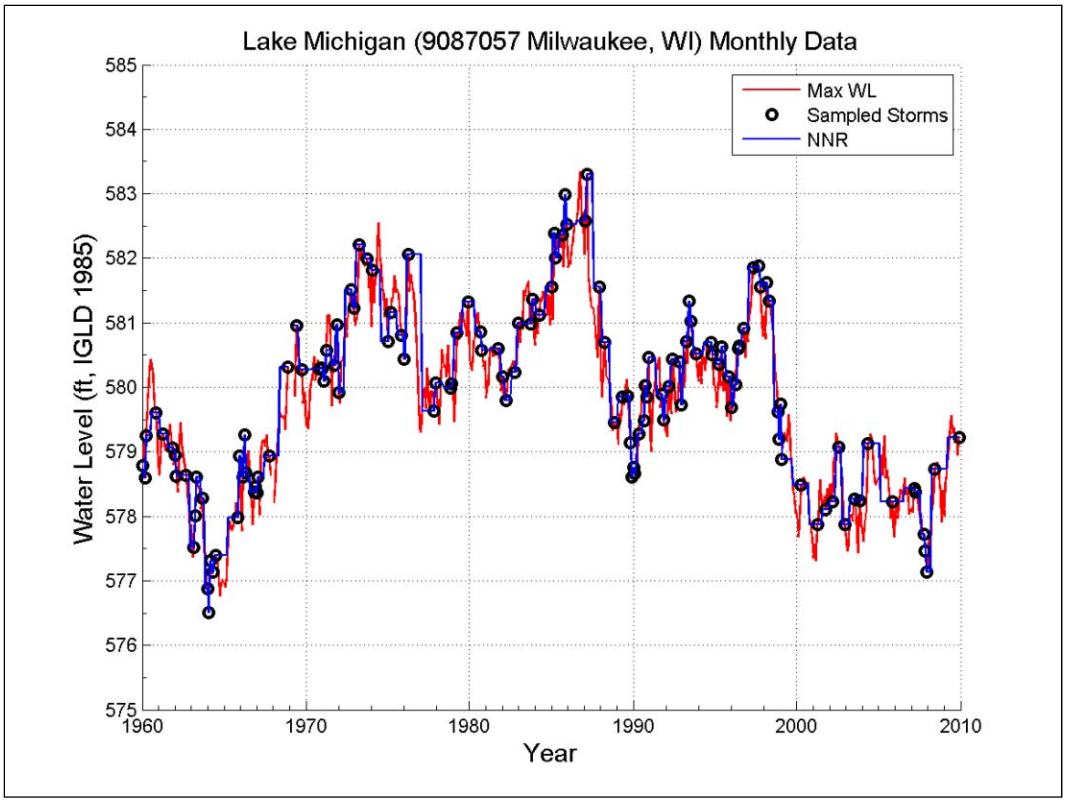



Figure 31. Distribution of 150-storm CSS across lake levels for Milwaukee, WI.

In all four Figures (28 through 31), the skill of the sampling is evident with all three high water level periods (i.e., mid 1970s, mid 1980s, and late 1990s) and all low water periods (mid 1960s, late 1970s, early 1990s and 2000s) captured. From Figures 30 and 31, it can be seen that the 150-storm CSS not only captures the decadal variation in water levels but also much of the higher frequency variation in water levels.

# 9 Evaluation of Bias in Water Level Measurements

As part of this study, the plausibility of bias in water level measurements was evaluated. The bias, or lack thereof, in measured hourly water levels was resolved and discussed in Melby et al. (2012). That analysis focused on comparing hourly water level data with 6-minute data. It was found that the bias in the hourly data was minimal. The hourly data bias, in terms of RMS deviation, was found to range between 0.10 to 0.20 ft (i.e., underestimation compared to the 6-minute data), on average.

The current task focused on assessing possible bias in monthly water level measurements. This issue is particularly relevant where continuous hourly water level data are not available. For this purpose, monthly water level data were compared to the full set of hourly water level measurements corresponding to the 1960-2010 period. Monthly water level statistics including monthly mean, maxima and minima are readily available from NOAA-GLERL. Storm surge was estimated from monthly data as the difference between monthly maxima and monthly mean, thus resulting in a single surge value for each month, here named surge monthly maxima. For comparison purposes, monthly maxima surges were also derived from hourly surge estimates by selecting the maximum surge value for each month. These hourly surge values are estimated as the difference between NOAA's verified tide and predicted tide.

Surge monthly maxima derived from both monthly and hourly data sources are displayed in Figures 32 and 33 for the Ludington and Milwaukee gages, respectively. In these two figures there are small discrepancies between the monthly and the hourly surge values but these seem to be random and do not constitute a bias. The largest observed errors are 40-50 percent, but these are isolated occurrences.

Scatter plots of surge monthly maxima from monthly data versus monthly maxima from hourly data were generated for each of the nine NOAA gages; the plots displayed in Figures 34 and 35, correspond to the Ludington and Milwaukee gages, respectively. Bias was also computed for every water level gage.

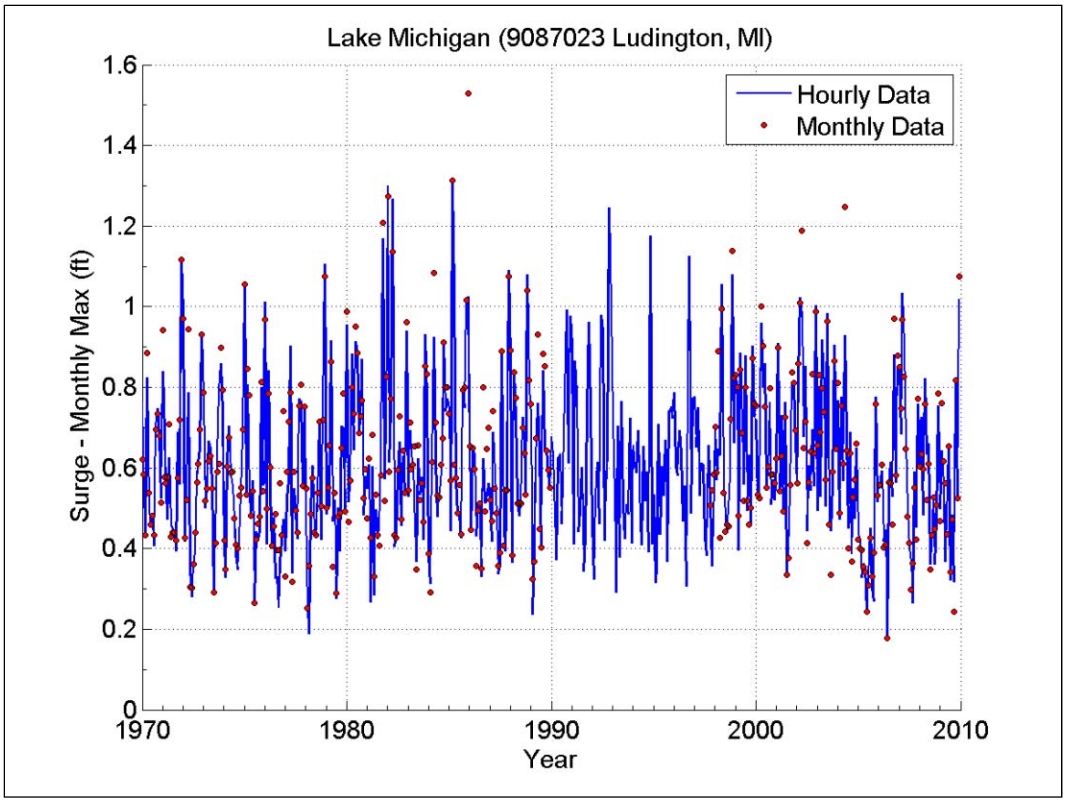



Figure 32. Surge monthly maxima from hourly and monthly data for Ludington, MI.

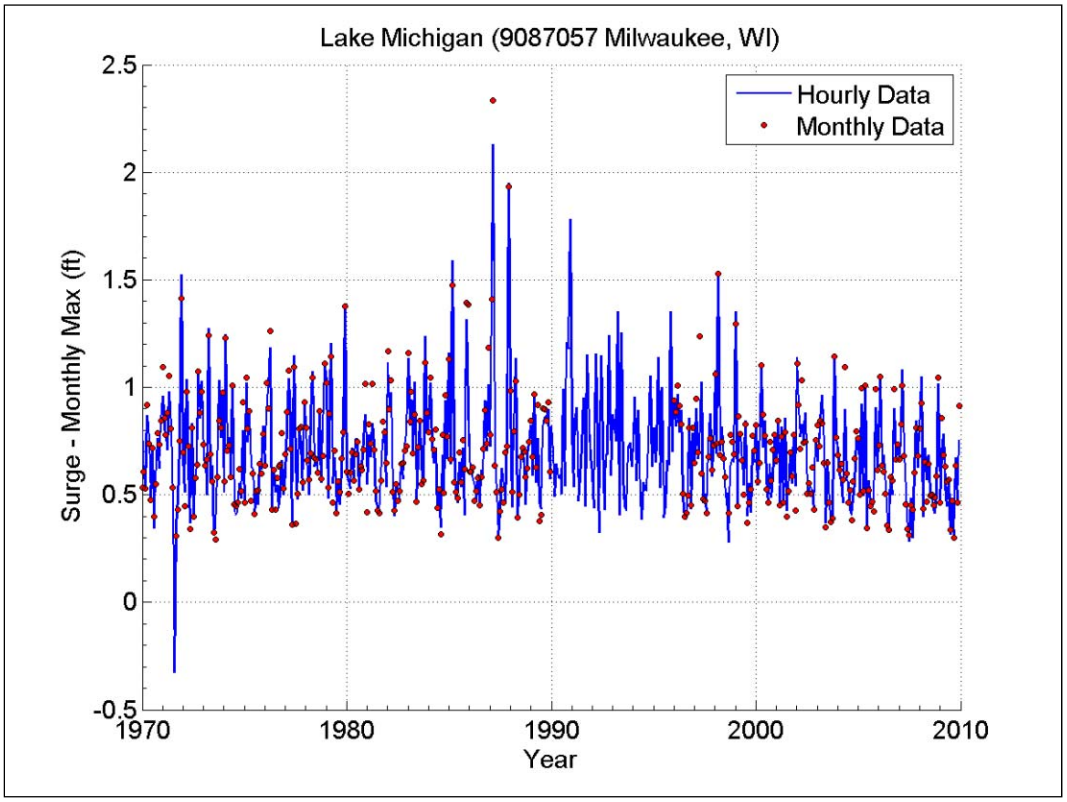



Figure 33. Surge monthly maxima from hourly and monthly data for Milwaukee, WI.

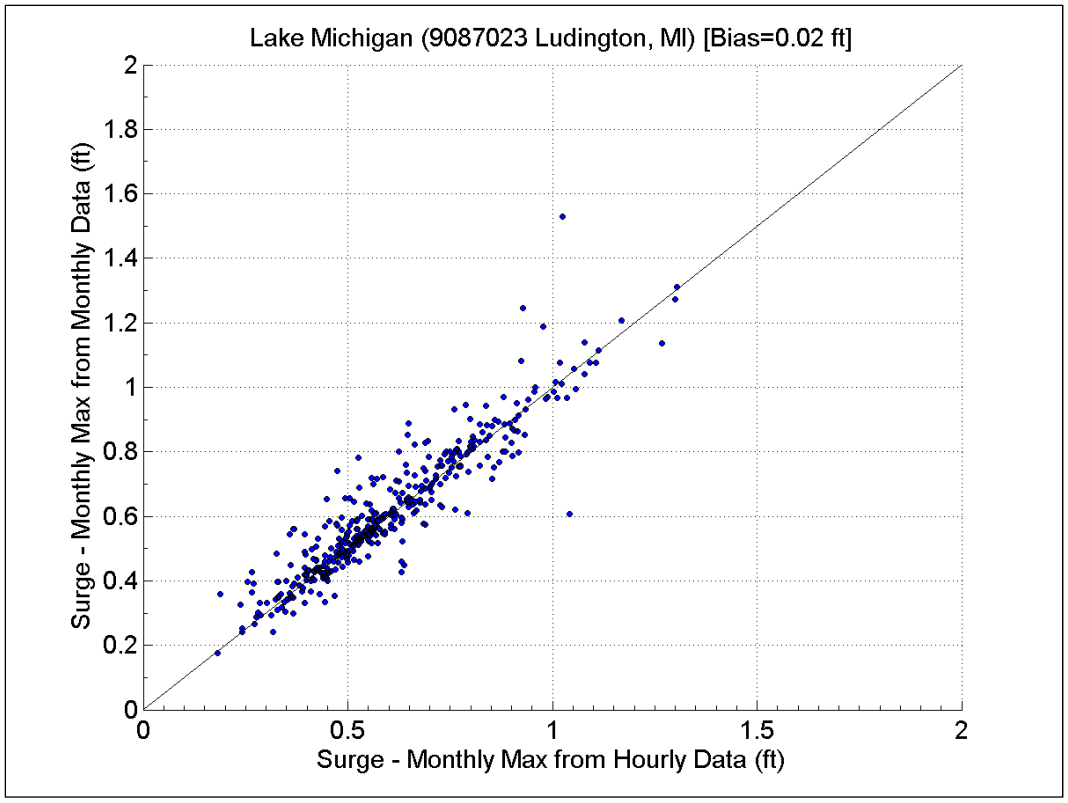



Figure 34. Surge monthly maxima from hourly and monthly data for Ludington, MI.

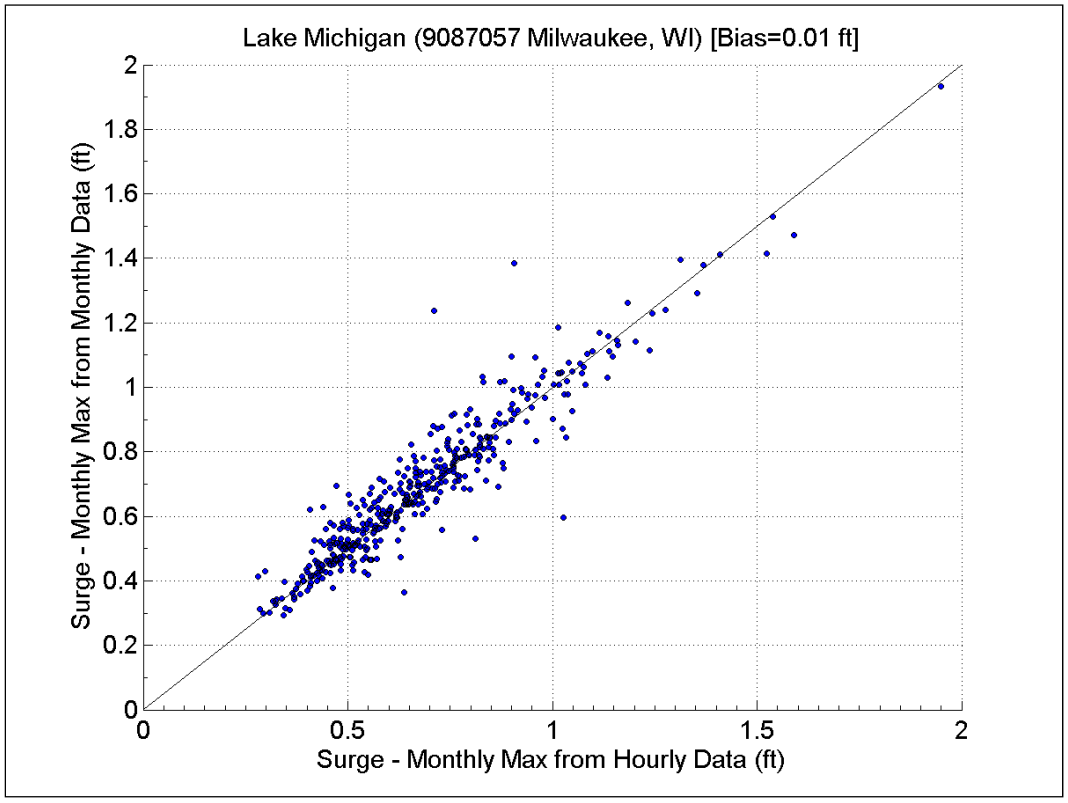



Figure 35. Surge monthly maxima from hourly and monthly data for Milwaukee, WI.

In the cases of Ludington and Milwaukee, the bias was estimated to be just 0.02 ft and 0.01 ft, respectively. These values were typical among all stations, with the Port Inland gage exhibiting the largest overall bias of 0.07 ft. Biases were also computed by decades (1970-79, 1980-89, 1990-99, and 2000-09) and also based on surge magnitude (0-1 ft, 1-2 ft, etc.). No bias of statistical significance was found.

A storm-by-storm assessment, however, showed that for some storms, surges derived from monthly data (i.e., difference between monthly maxima and monthly mean) can significantly overestimate surge values when compared to those computed from hourly data. This occurs mainly due to the stair-step effect of the monthly data (both maxima and mean). When intra-monthly water level is trending, often, the difference between monthly maxima and monthly mean will just reflect changes in mean lake level from month to month, thus overstating or understating surge magnitudes and falsely identifying events.

Therefore, surge events identified using monthly mean water level data should be considered only as preliminary candidates when constructing a CSS. If pre-1970s continuous hourly water level data are not available, pre-1970s surge events, based on monthly data, can be cross-referenced with measured wind data (wind is a primary forcing mechanism for storm surge and waves), wave data, surrogate wave estimates, or WIS hindcasts, to validate their authenticity. Rejecting an entire decade from consideration, such as the 1960s would result in a reduction of the record length to just 40 years which could significantly increase the error in water level probability distributions and extreme statistics, and further increase the uncertainty in extrapolation to estimate one percent and 0.2 percent annual chance water levels.

# 10 Still Water Level – Full Storm Set

The full storm set (FSS) is defined as a set of storms where all events are sampled from the same spatial location. In the case of Lake Michigan, nine FSS were generated, one at each of the nine NOAA water level gage locations. Each FSS represents the complete set of unique storms that have affected each individual location.

In the case of SWL, the FSS were generated using the entire record length available at each gage location. Thus, for SWL these lengths ranged from 38 years (Kewaunee, WI) to 107 years (Calumet Harbor, IL).

As previously discussed, TWL (which includes wave effects) is the reference level used to establish BFEs. In the case of TWL, we are limited to the 1960-2010 period where waves and wave runup can be estimated. However, analyzing SWL data and computing SWL statistics allow assessment of various aspects of the methodology in light of having longer periods of record; even 100+ years of record, depending on the station. Having more years of water level data allows for a more robust analysis of the storm sampling technique and storm sample size decisions, and it enables validation of the adopted statistical approach.

The process to develop each of the SWL-FSS is summarized as follows:

1. For each water level gage, select the most significant storms using the POT sampling technique on the storm surge data.
2. A preliminary sample intensity (lambda) of approximately 4- 5 storms per year should be used, based on previously reported results in Melby et al. (2012). This value of lambda results in a preliminarily large sample population of extreme events. In the case of Lake Michigan, a lambda value of 4 would yield anywhere from 150- 430 total storms per gage.
3. A lower-bound return period threshold ($RP_{th}$) corresponding to high-frequency events with return periods of 0.1-0.9 years is then used to improve the GPD fits. The process of determining $RP_{th}$ is iterative and the goodness of the fit is evaluated using CDF plots and Q-Q plots.
4. This threshold can result in a reduced number of storms included in the distribution, equivalent to a lambda of 1-3 storms per year.
5. Fit distributions using the Q-Q optimization approach.

The purpose of this approach is to obtain better Generalized Pareto Distribution (GPD) fits in the high-frequency portion of the distributions and return period plots. Preliminary assessments showed that the incorporation of a lower-bound $RP_{th}$ corresponding to a return period of roughly one year resulted in better high-frequency tail fittings and overall better distribution shapes. Further analysis determined that this lower-bound threshold varies from location to location; in the case of Lake Michigan gages, this threshold ranges from 0.1 to 0.9 years, with typical values around 0.5-0.8 years. The goodness of these new fittings was evaluated using the cumulative distribution function (CDF) and quantile-quantile (Q-Q) plots.

The Q-Q plots in this report display the quantiles of the actual water level data (either SWL or TWL) vs. the theoretical quantiles of the GPD that was fitted to that data. Superimposed on the Q-Q plots are a robust linear fit of the Q-Q data and a 45-degree line to help evaluate linearity. The robust linear fit is deemed so because it is actually a line that joins the first quartile (Q1) and the third quartile (Q3) of, both, the water level data and the GPD fit. This segment between Q1 and Q3 is known as the interquartile range (IQR = Q3 – Q1) and it is defined as a robust order statistic.

In theory, the closer the robust Q-Q linear fit (red dashed line shown in Q-Q plots) is to the 45-degree line, the better the GPD fit. A metric was established to interactively identify the lower-bound threshold $RP_{th}$ that minimizes the difference between the slopes of both of these lines. Since the slope of the 45-degree line is equal to unity, the absolute difference between both slopes is expressed as follows:

$$\varepsilon = \left| \left( \frac{\Delta y}{\Delta x} \right)_{Q-Q} - 1 \right| \tag{24}$$

where

$\varepsilon$ = absolute value of the difference between the robust Q-Q linear fit and the 45-degree reference line, and

$$\left( \frac{\Delta y}{\Delta x} \right)_{Q-Q} = \frac{(Q_3 - Q_1)_{\text{observed data}}}{(Q_3 - Q_1)_{\text{GPD fit}}} \tag{25}$$

is the slope of the robust Q-Q linear fit.

The value of $\varepsilon$ is evaluated iteratively using several $RP_{th}$ values until the minimal value of $\varepsilon$ is identified. The CDF plots comparing the GPD CDF to the empirical CDF are used for qualitative assessment of the distribution fits.

Different root-mean-square (RSM) values are evaluated as additional metrics. For example, RMS1, is defined as the root-mean-square deviation of all events with return period equal to or greater than 1 year; likewise, RMS10, accounts for all events with return period equal to or greater than 10 years. RMS1 shows how good the overall fits and the shapes of the distributions are. RMS10 focuses on the upper (extreme) tail of the distributions. One caveat related to these RMS values is that occasionally RMS1 can be decreasing while RMS10 is increasing, and vice versa. The Q-Q optimization approach will result in adequate balancing of both RMS1 and RMS10.

Q-Q plots can be generated for visualization purposes, but they are not required as part of the Q-Q optimization approach. The methodology for the Q-Q optimization of extreme water level distributions is summarized as follows:

1. Fit GPD to extreme water levels (i.e., SWL, TWL) using a lambda of 4-5 storm events per year, to have a large number of observations.
2. Compute the Q1 and Q3 values for both the observed extreme water levels and the GPD fit.
3. Compute the slope of the robust Q-Q linear fit, as previously defined.
4. Determine the lower-bound return period thresholds, $RP_{th}$. From the GPD, compute the water levels corresponding to return periods ranging from 0.01-year to 1.0-years; a 0.01-year interval was used for this study for a total of 100 different $RP_{th}$.
5. Use each of the $RP_{th}$ as a new threshold to redo the GPD fitting.
6. For each $RP_{th}$/GPD fit compute the slope of the robust Q-Q linear fit.
7. For each $RP_{th}$/GPD fit, compute ε, the absolute difference between the slope of the robust Q-Q linear fit and the 45-degree reference line (slope = 1).
8. Finally, select the $RP_{th}$/GPD combination corresponding to the lowest value of $\varepsilon$ as the optimized fit.

This optimization process effectively reduces the value of lambda used for the GPD fittings. Strictly considering SWL-FSS values in the cases of Ludington and Milwaukee, from an initial sample intensity of four, lambda was reduced to 2.9 and 1.8 events per year, respectively. Further

interpretation of this analysis seems to indicate that the Ludington location was affected by roughly 145 extreme SWL events during the 1960-2010 period; while the Milwaukee location was affected by approximately 90 extreme SWL events. At least these are the numbers of extreme events that belong to each respective extreme water level distribution.

Figures 36 -47 show SWL-FSS Q-Q plots, CDF plots, and RP plots corresponding to the Ludington and Milwaukee water level gages. Plots are shown for both locations, including and excluding convective storm events.

Table 5 shows the SWL-FSS estimates corresponding to one percent and 0.2 percent annual chance SWL values, both with and without convective storms, at each of the nine NOAA water level gage locations (location names are listed in Table 2). The influence of neglecting summertime convective events is generally small and results in differences of less than 0.2 ft, with the greatest differences observed at the Sturgeon Bay, WI site (9087072) and at Green Bay, WI (9087079), where differences are as much as 0.3 to 0.4 ft but no significant lake-wide bias is evident. Note that only the last five digits of each station ID are shown on the table (i.e., 87079 instead of 9087079).

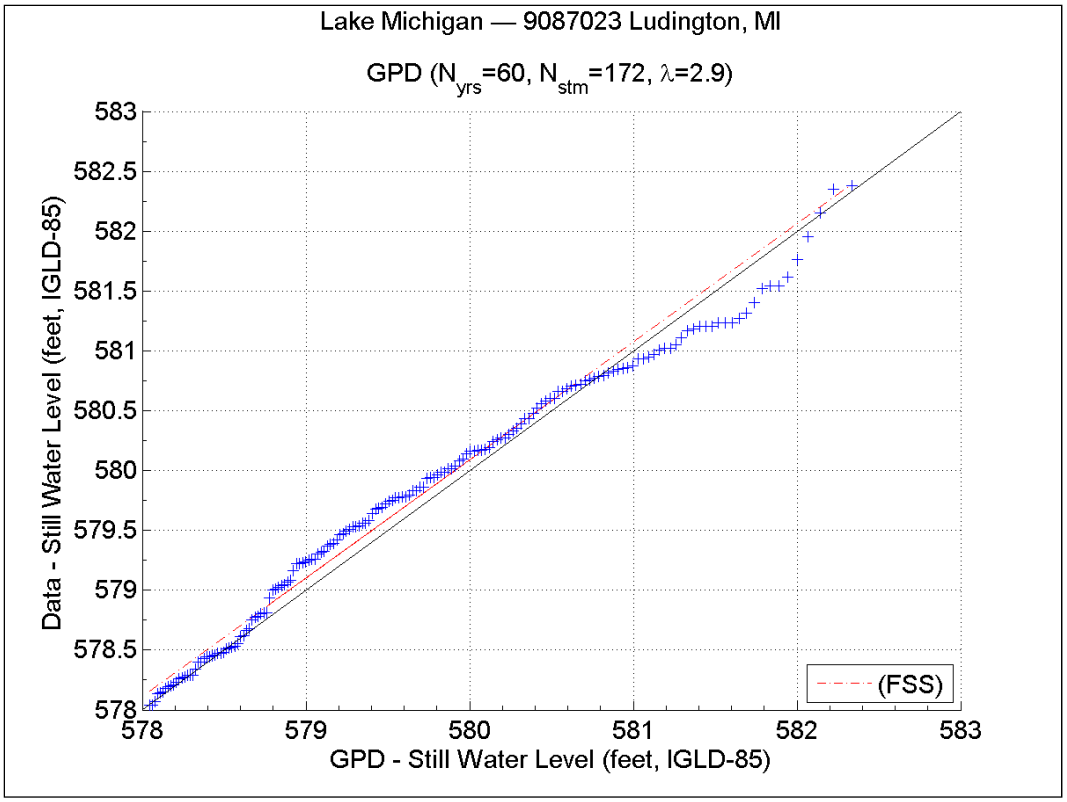



Figure 36. SWL Q-Q plot from FSS (with convective storms) for Ludington, MI.

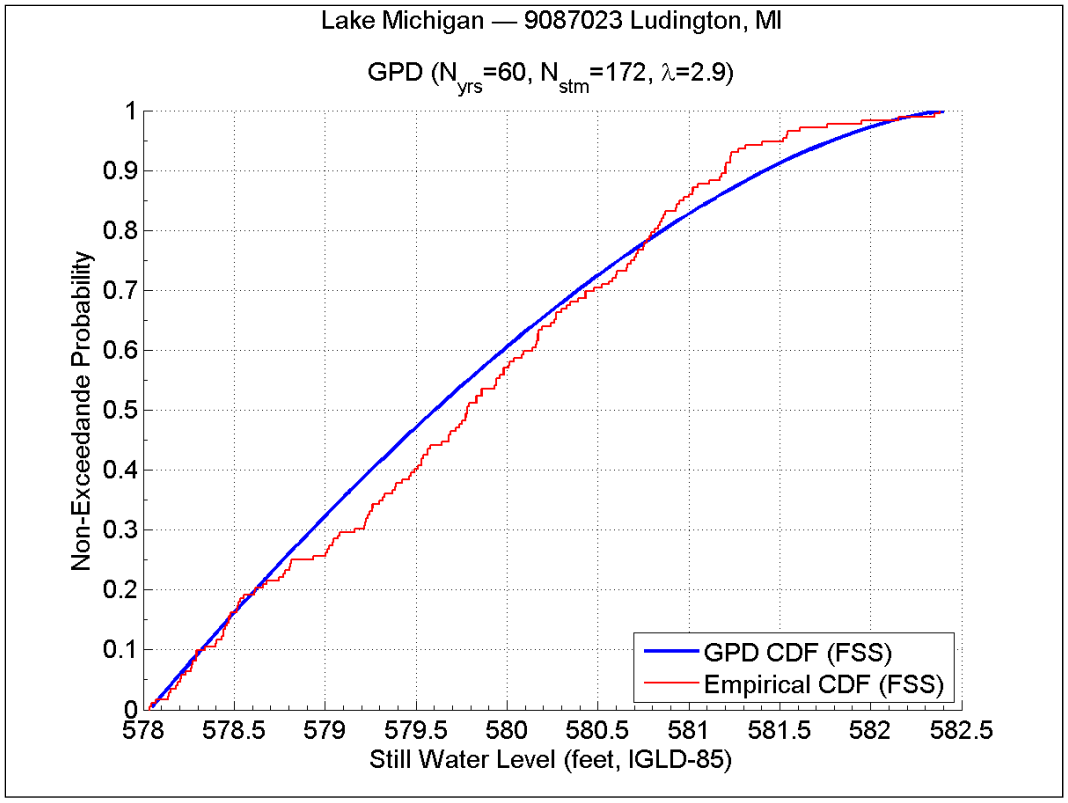



Figure 37. SWL CDF plot from FSS (with convective storms) for Ludington, MI.

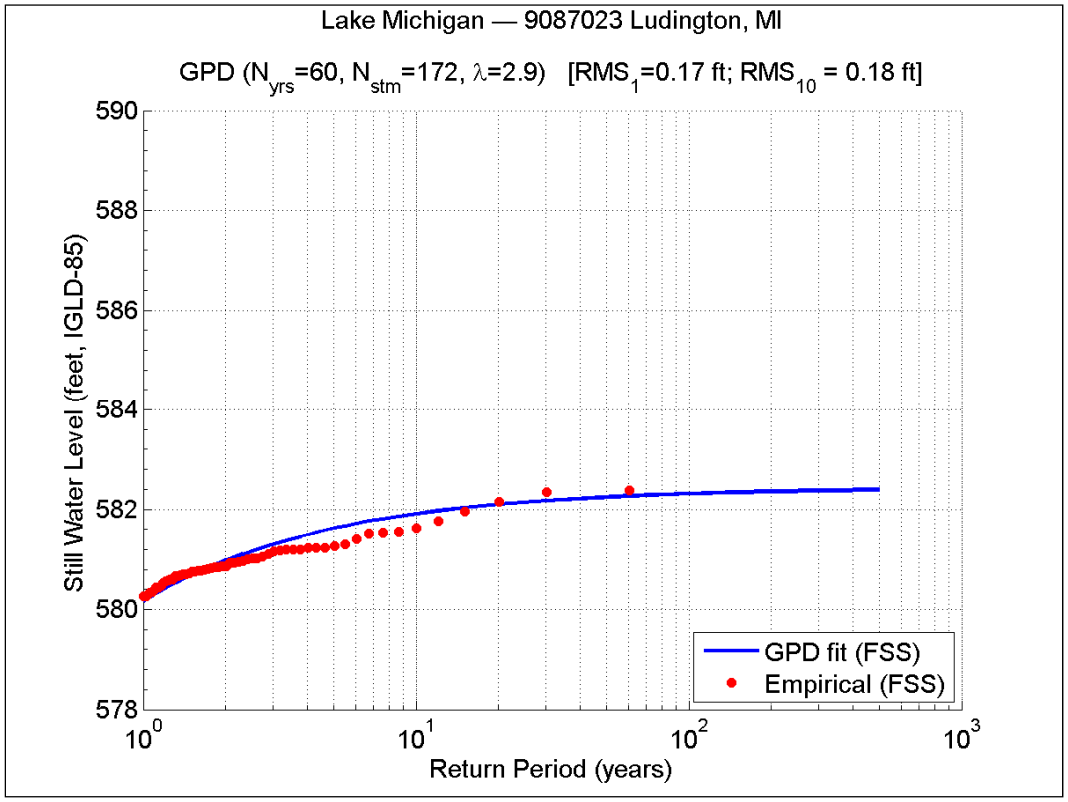



Figure 38. SWL RP plot from FSS (with convective storms) for Ludington, MI.

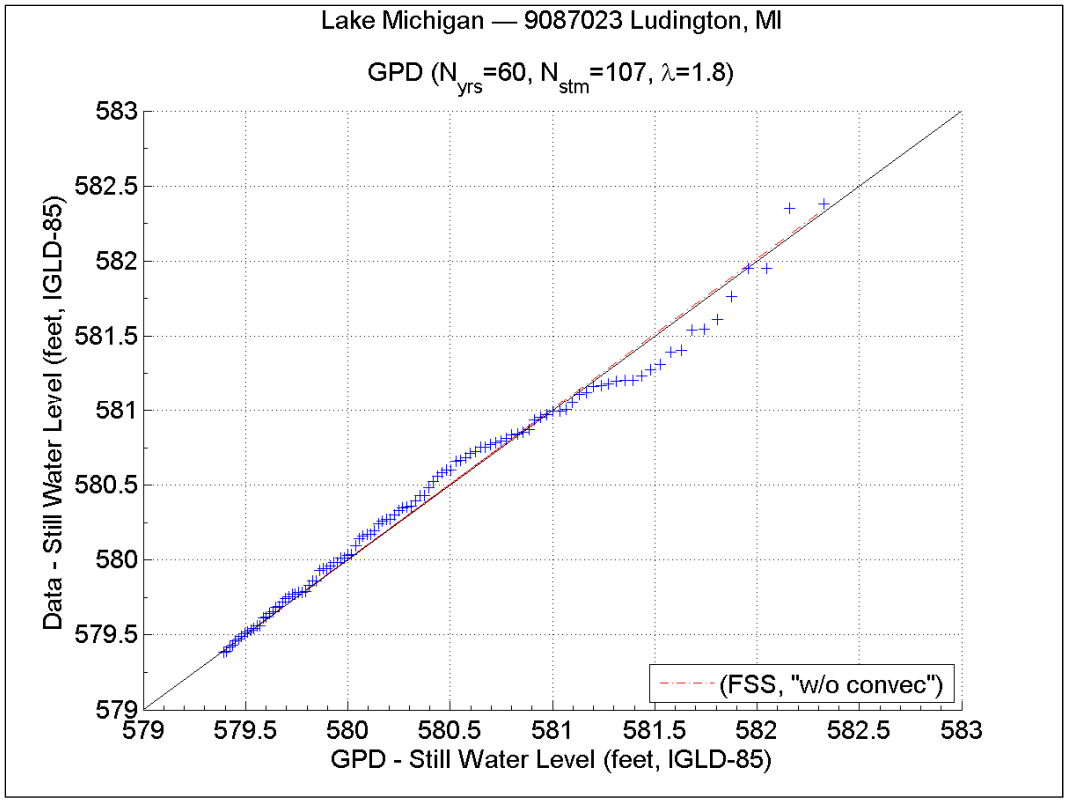



Figure 39. SWL Q-Q plot from FSS (no convective storms) for Ludington, MI.

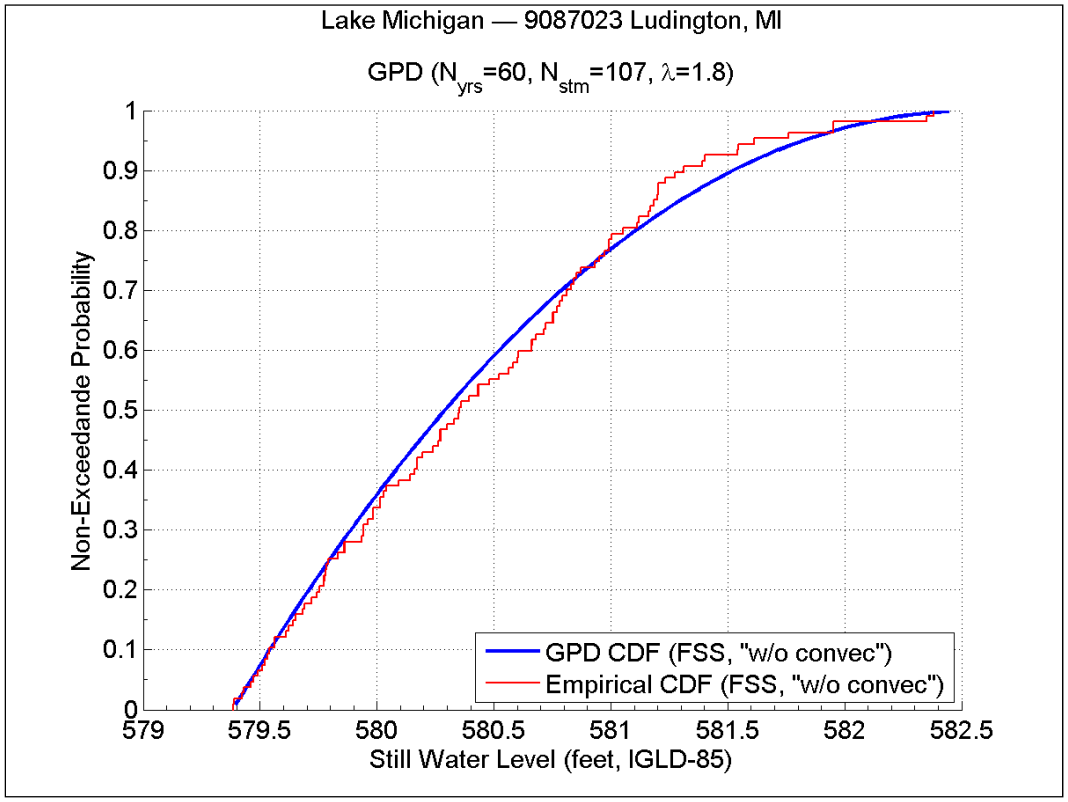



Figure 40. SWL CDF plot from FSS (no convective storms) for Ludington, MI.

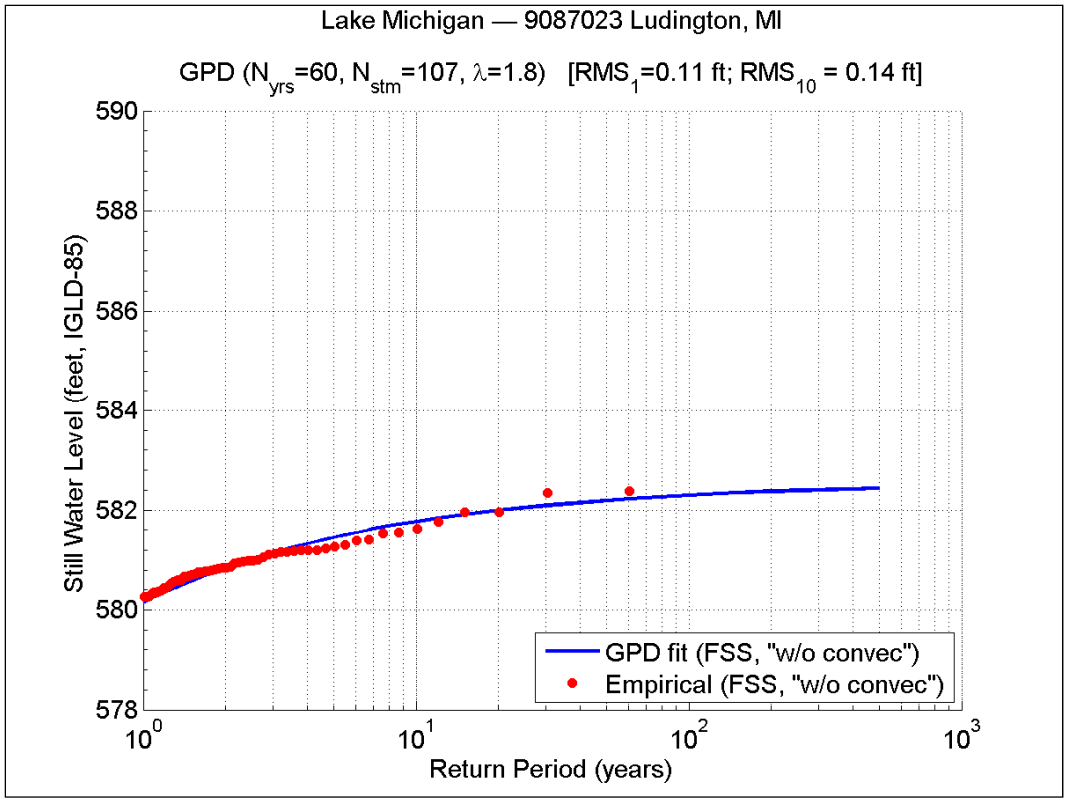



Figure 41. SWL RP plot from FSS (no convective storms) for Ludington, MI.

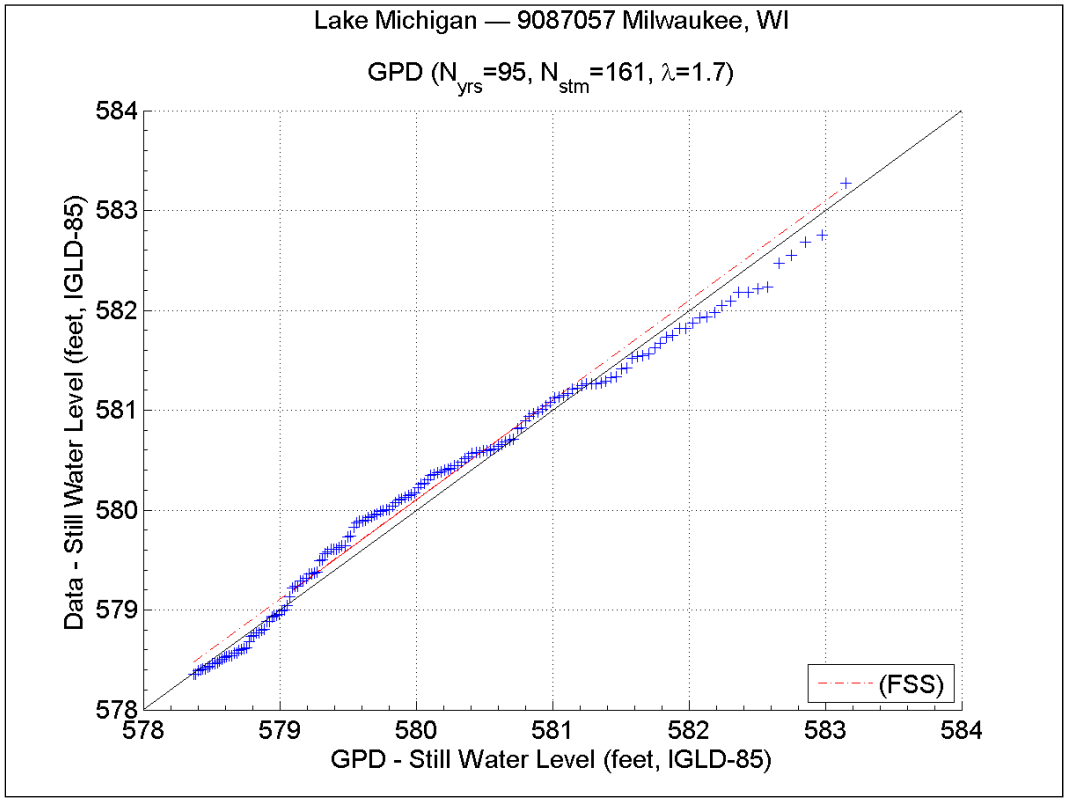



Figure 42. SWL Q-Q plot from FSS (with convective storms) for Milwaukee, WI.

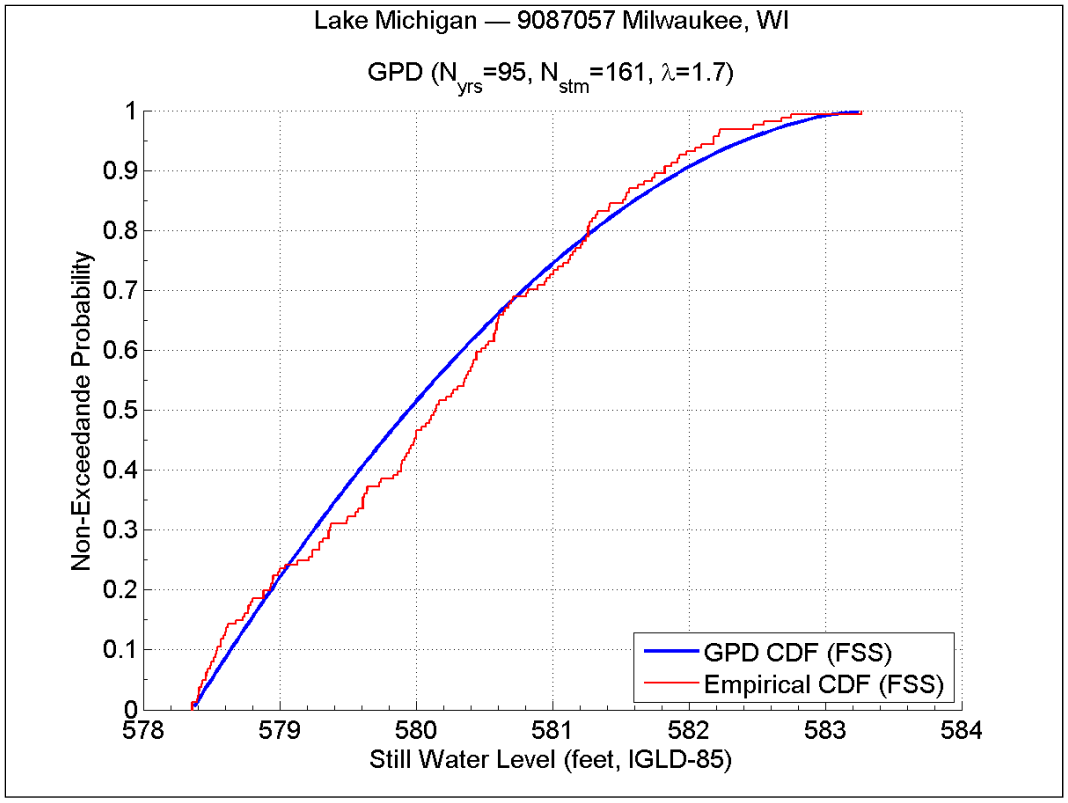



Figure 43. SWL CDF plot from FSS (with convective storms) for Milwaukee, WI.

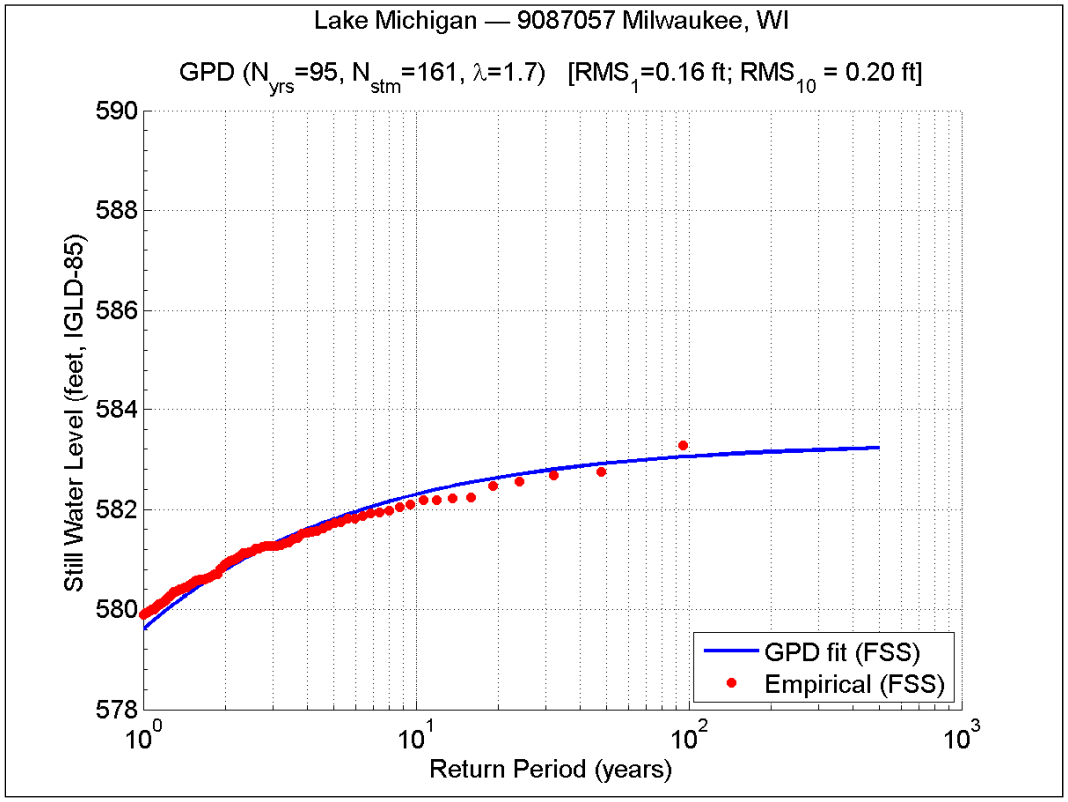



Figure 44. SWL RP plot from FSS (with convective storms) for Milwaukee, WI.

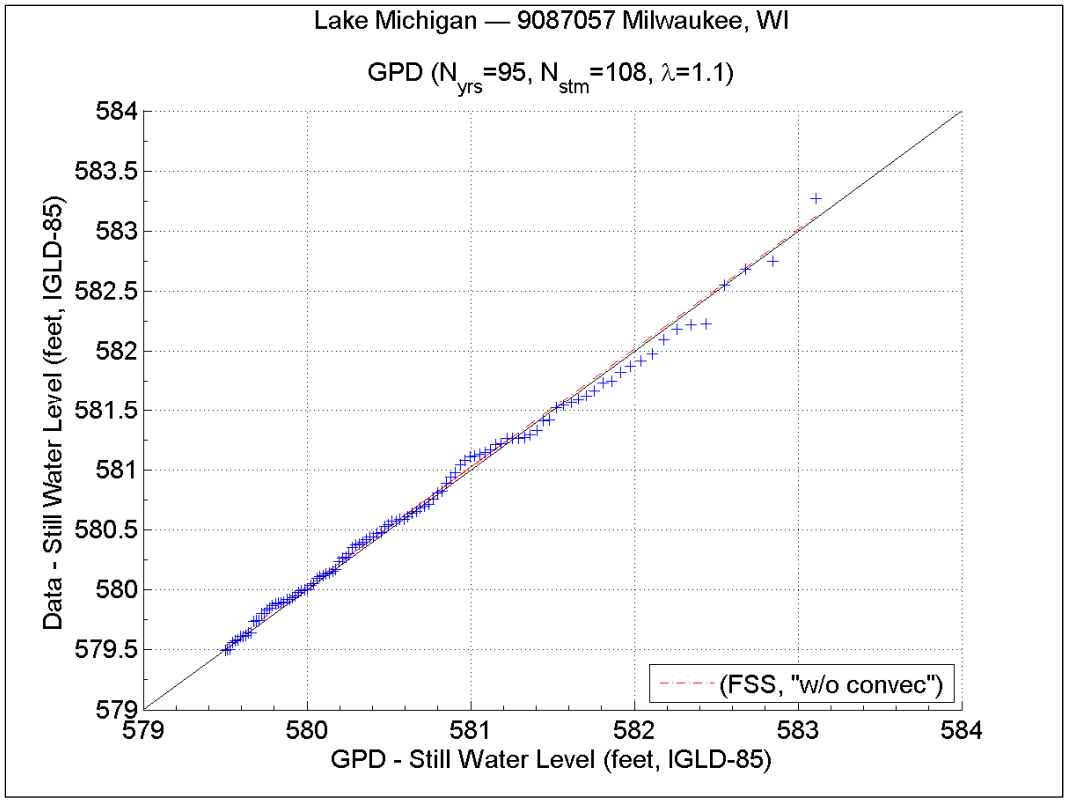



Figure 45. SWL Q-Q plot from FSS (no convective storms) for Milwaukee, WI.

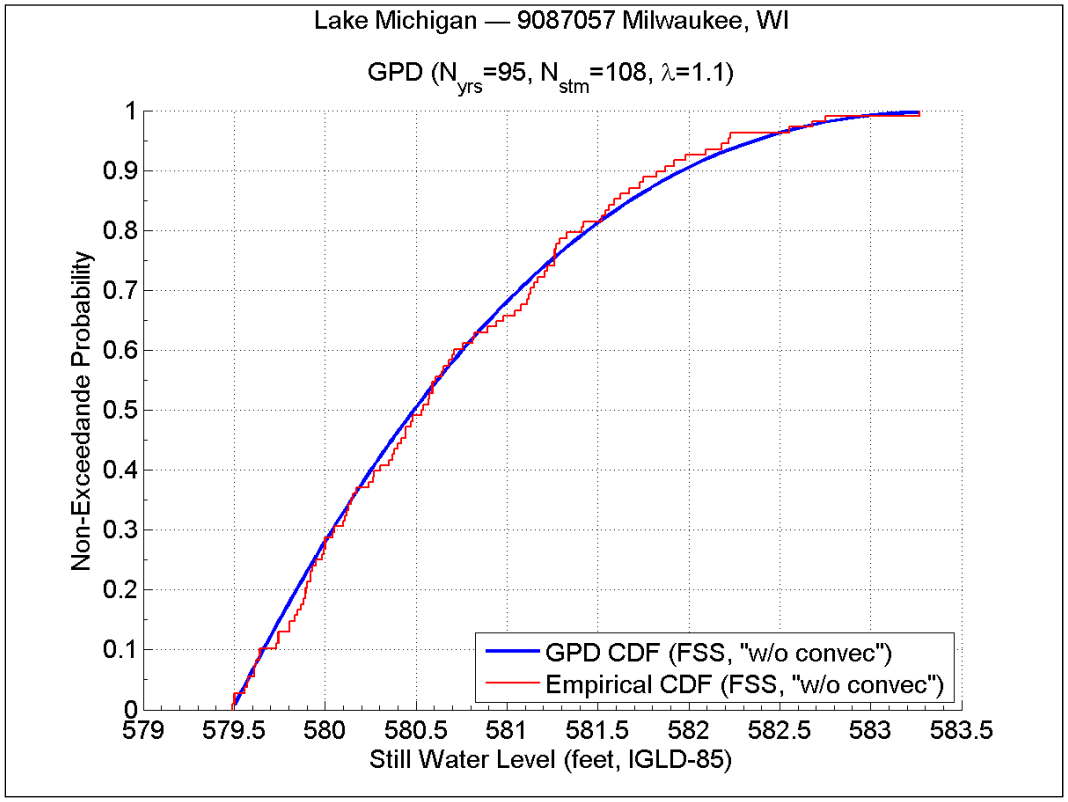



Figure 46. SWL CDF plot from FSS (no convective storms) for Milwaukee, WI.

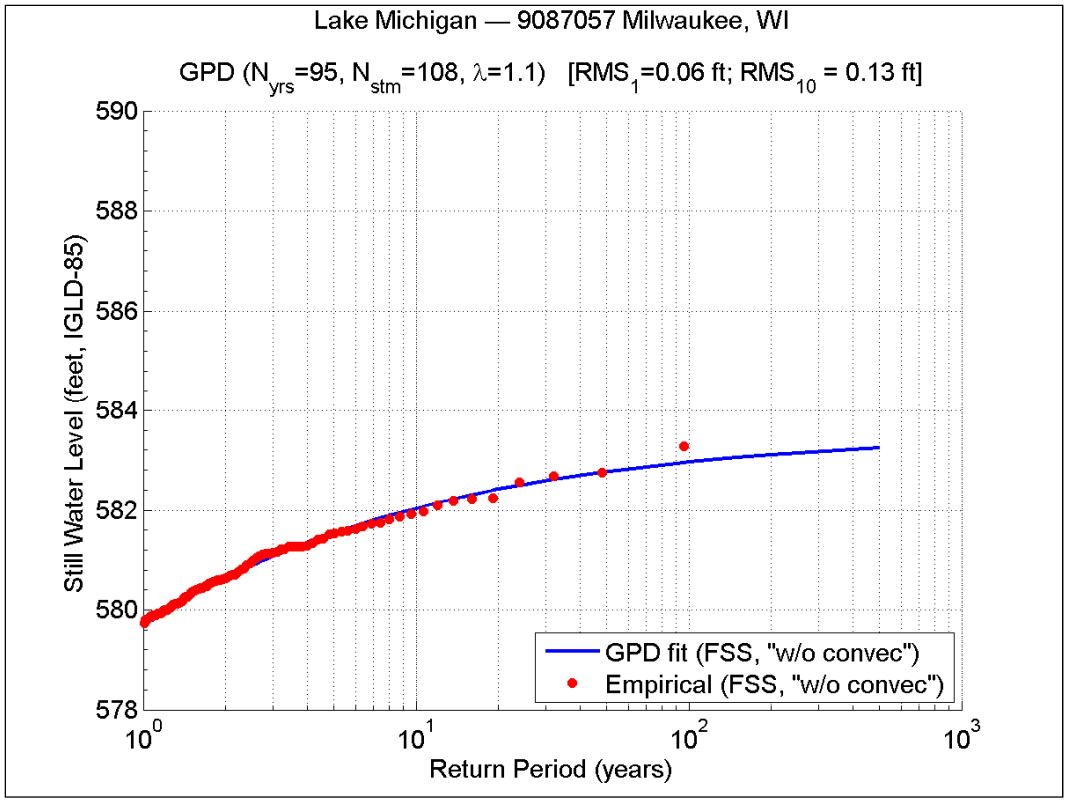



Figure 47. SWL RP plot from FSS (no convective storms) for Milwaukee, WI.

Table 5. SWL-FSS values for one percent and 0.2 percent annual chance events.

| SWL-FSS with summer events (ft, IGLD-85) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Prob (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 582.40 | 583.29 | 583.67 | 583.24 | 583.20 | 583.02 | 584.17 | 585.24 | 583.17 |
| 1 | 582.32 | 582.82 | 583.46 | 583.07 | 582.93 | 582.87 | 584.05 | 584.18 | 582.88 |
| SWL-FSS without summer events (ft, IGLD-85) | | | | | | | | | |
| Prob (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 582.45 | 583.07 | 583.65 | 583.26 | 582.98 | 582.74 | 584.61 | 585.41 | 583.13 |
| 1 | 582.31 | 582.93 | 583.44 | 582.97 | 582.76 | 582.57 | 584.22 | 584.23 | 582.79 |

# 11 Still Water Level – Composite Storm Set

A composite storm set (CSS) is defined as a single set, or composite, of storms derived from storms sampled from different gage locations around a given lake. For the 50 year period, between 1960 and 2010, as discussed in previous sections, it was determined that the necessary number of storms to generate the CSS was 150. The process to generate SWL-CSS and the corresponding SWL distributions can be outlined as follows:

1. For a 150-storm CSS, select a large number of storms from each water level gage (e.g., 30 storms per gage). This will result in sampling the required number of extreme events plus enough additional storms to replace duplicates.
2. Construct a preliminary CSS using the top *N* storms per gage, as:

$$N = ceiling\left(\frac{n_{storms}}{n_{gages}}\right) \tag{26}$$

where

$n_{storms}$ = optimal number of total storms
$n_{gages}$ = number of water level gages
$ceiling$ = round up function

For example, in the Lake Michigan case: $N = ceiling\ (150/9)$, which yields 17 storms/gage; for a total of 17 x 9 = 153 storms.

3. Identify, remove, and replace duplicate storms, which are storms that are part of the top *N* storms at more than one gage. When duplicate storms are identified, the storm with the highest ranking for its corresponding gage should be retained. Then, the same storm having a lower ranking at other locations should be removed from those sets and replaced with the next storm in line from the same gage location.
4. Repeat the replacement process until the desired set of 150 unique storms is obtained.

The following are the SWL-CSS Q-Q, CDF and RP plots corresponding to the Ludington and Milwaukee water level gages (Figures 48-59). Note that the Q-Q optimization results in a reduced number of storms being used to compute the statistics. In the case of Ludington, out of 150 CSS storms, roughly 70 storms are considered. This number varies from gage to gage.

Table 6 shows the SWL-CSS estimates corresponding to one percent and 0.2 percent annual chance SWL values, both, with and without convective storms. As was the case of the SWL-FSS results, differences between the with- and without-convective event sets are generally small, except at Sturgeon Bay. The Sturgeon Bay gage is located inside the Sturgeon Bay canal, and this area might be more sensitive to convective events than other areas of Lake Michigan.

## 11.1 Still water level: full storm set vs. composite storm set

Table 7 shows the comparison between SWL-CSS and SWL-FSS. Table 7.a represents the differences between SWL-CSS and SWL-FSS, for the one percent and 0.2 percent annual chance values, both including summer convective events. Table 7.b shows the differences between SWL-CSS without convective events and SWL-FSS with convective events, also for the one percent and 0.2 percent annual chance values. The SWL-CSS values without summer convective events are the SWL results from the statistical analysis and storm sampling approach, and they are compared to the SWL-FSS values with summer convective events because the latter represents the *true* SWL distribution.

It can be seen that, with the exception of three stations, differences in extreme SWL values are typically below 0.20 ft, with an average over-prediction of 0.04 ft using the 150-storm CSS. The relatively large discrepancies that are observed for gages 9087068 (Kewaunee, WI) and 9087096 (Port Inland, MI) are mainly due to their short record lengths. The Kewaunee and Port Inland gages have record lengths of 37 and 45 years, respectively. Therefore, it is uncertain whether the SWL-FSS distributions for both of these stations are representative of the true SWL distributions.

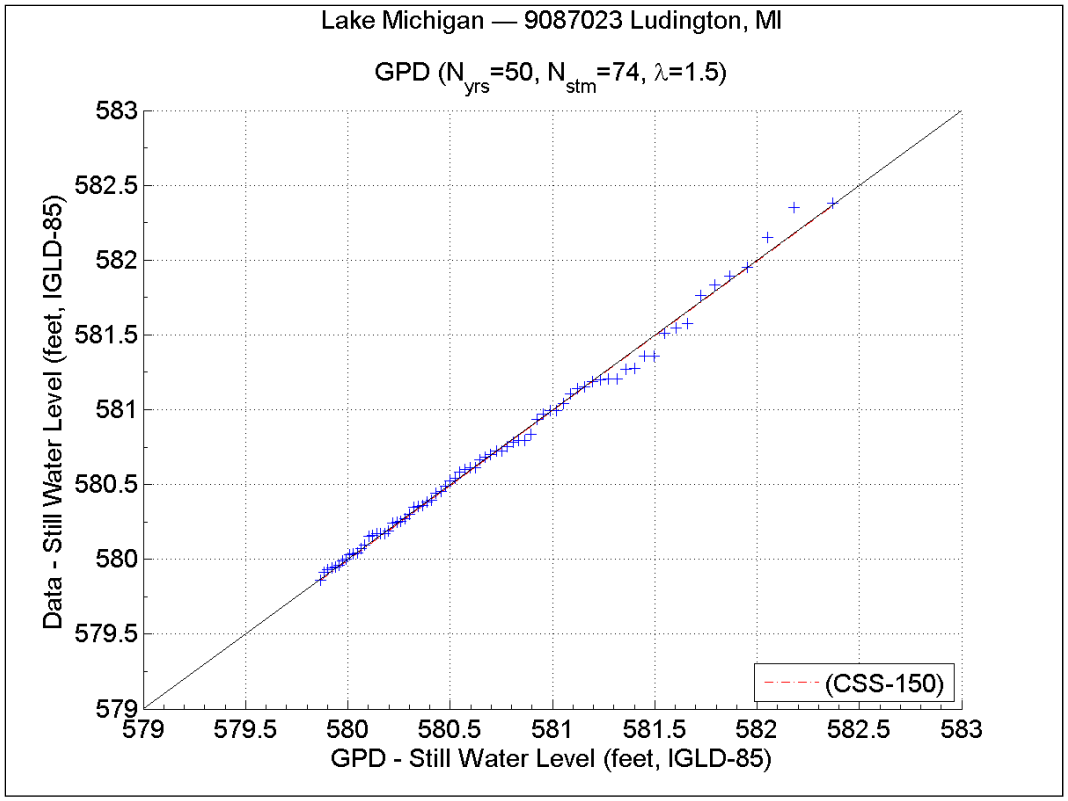



Figure 48. SWL Q-Q plot from CSS (with convective storms) for Ludington, MI.

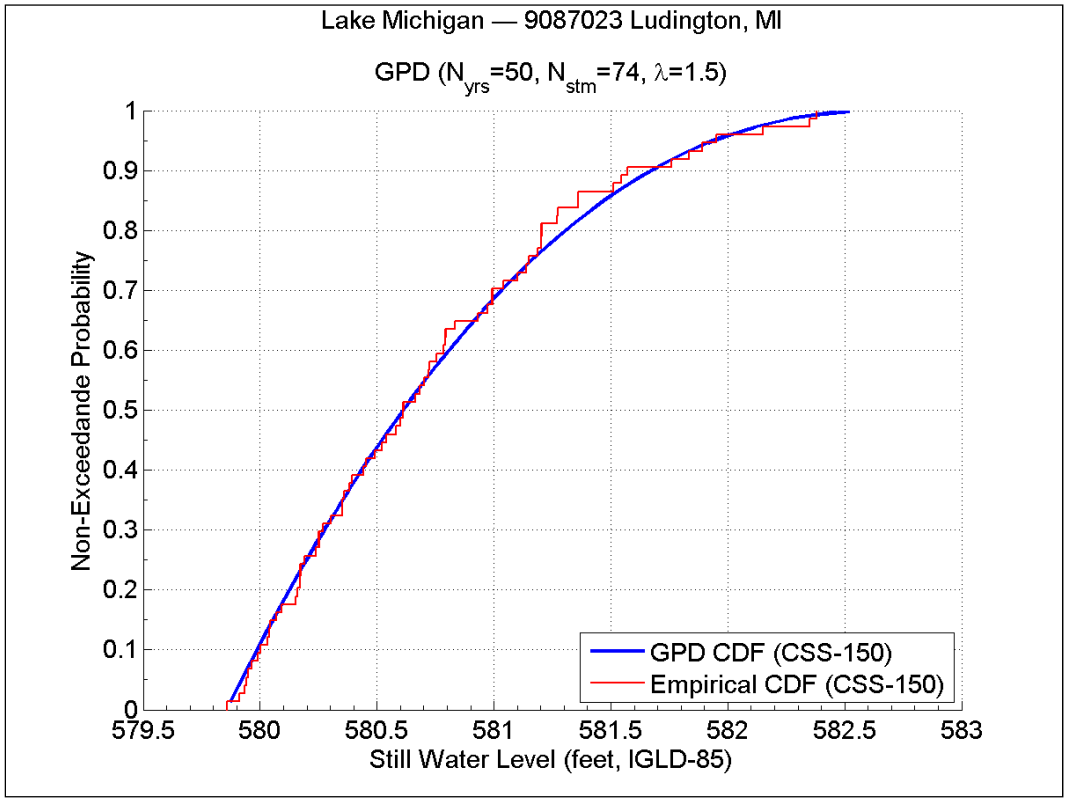



Figure 49. SWL CDF plot from CSS (with convective storms) for Ludington, MI.

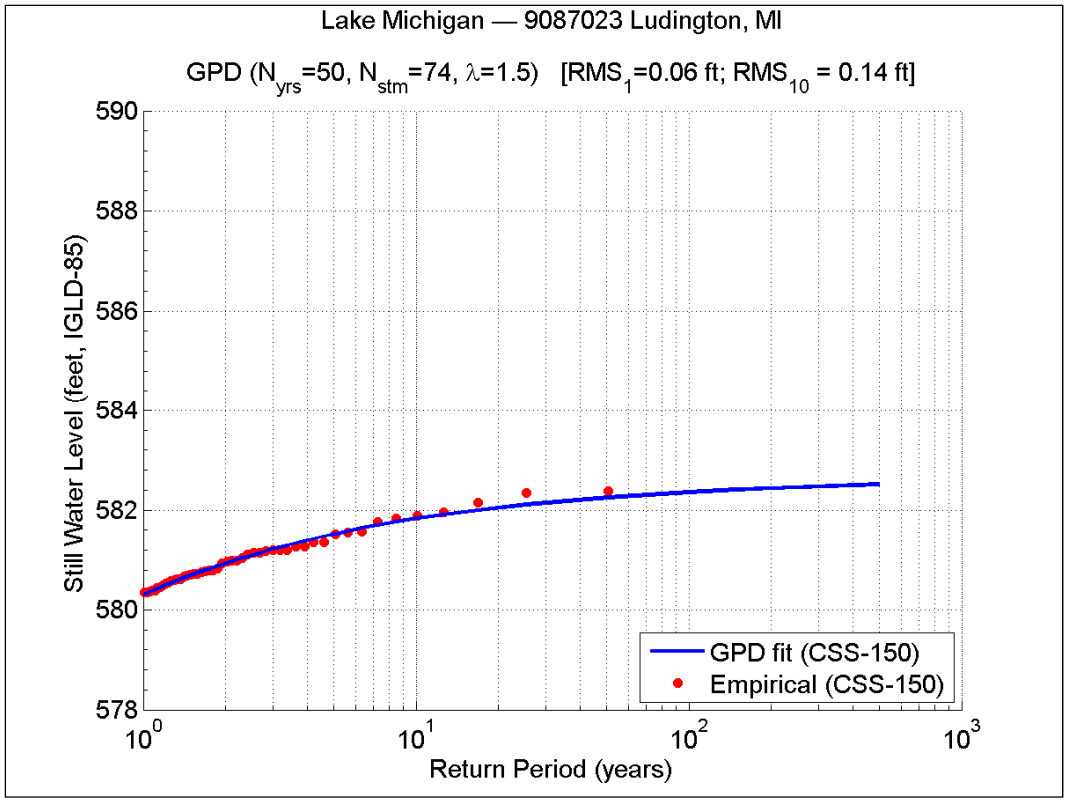



Figure 50. SWL RP plot from CSS (with convective storms) for Ludington, MI.

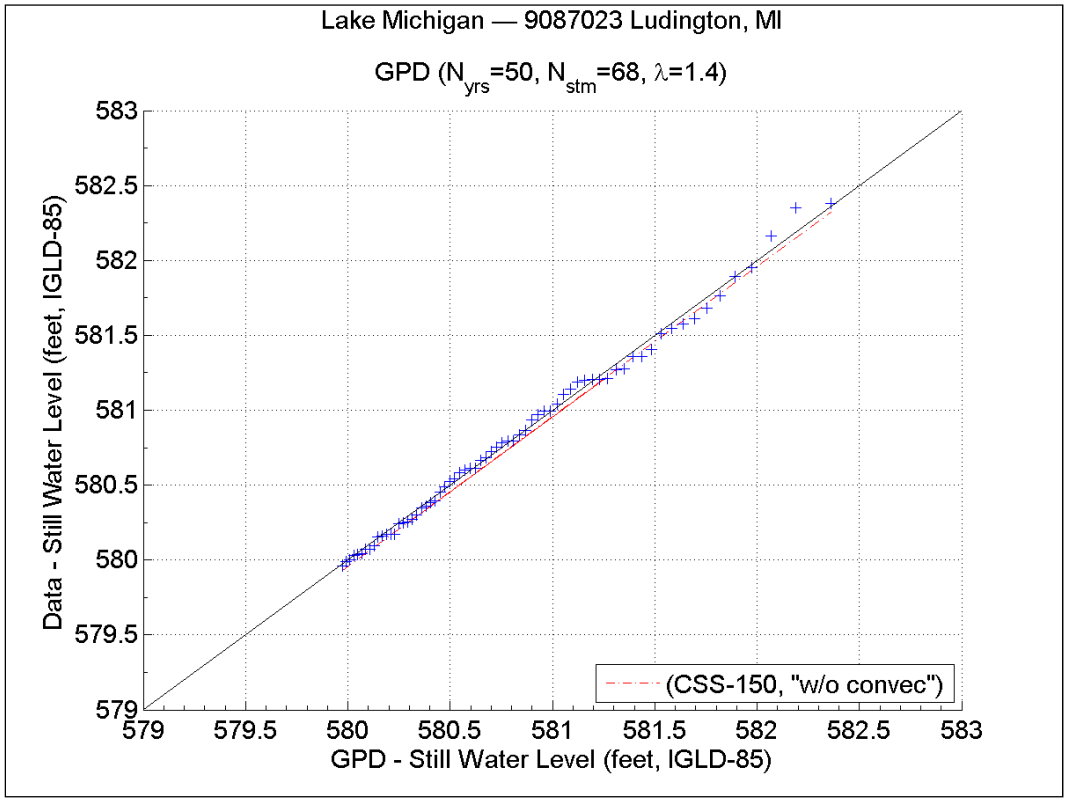



Figure 51. SWL Q-Q plot from CSS (no convective storms) for Ludington, MI.

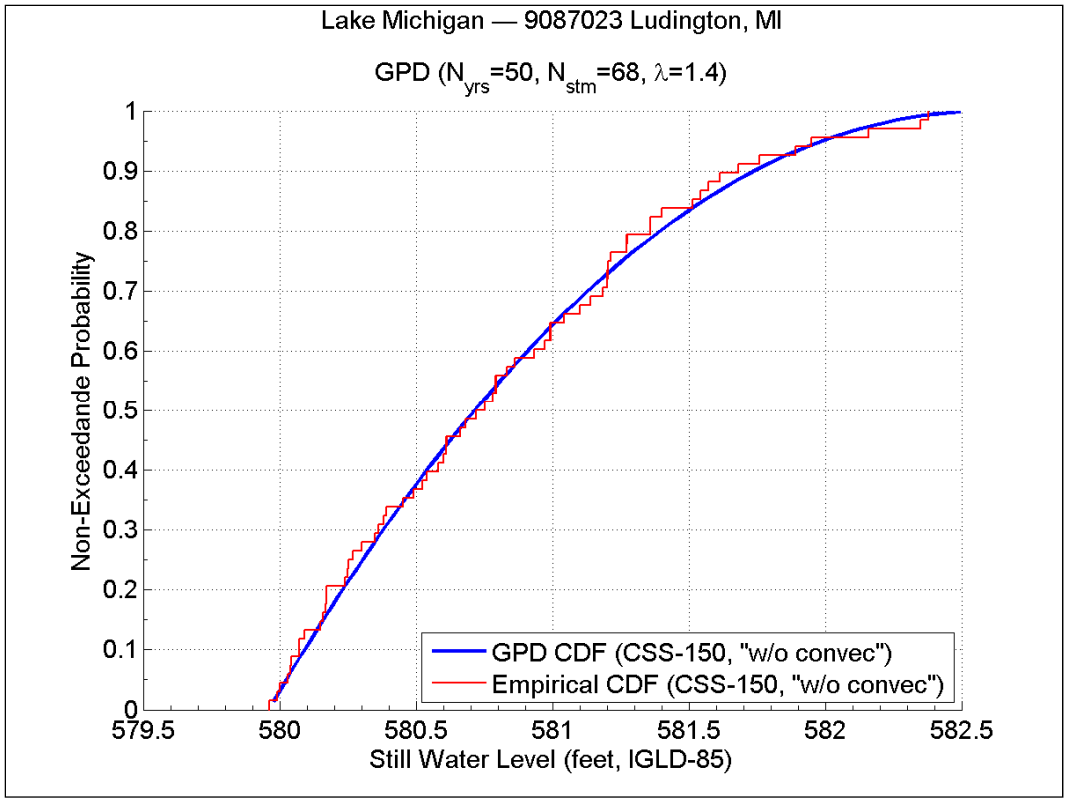



Figure 52. SWL CDF plot from CSS (no convective storms) for Ludington, MI.

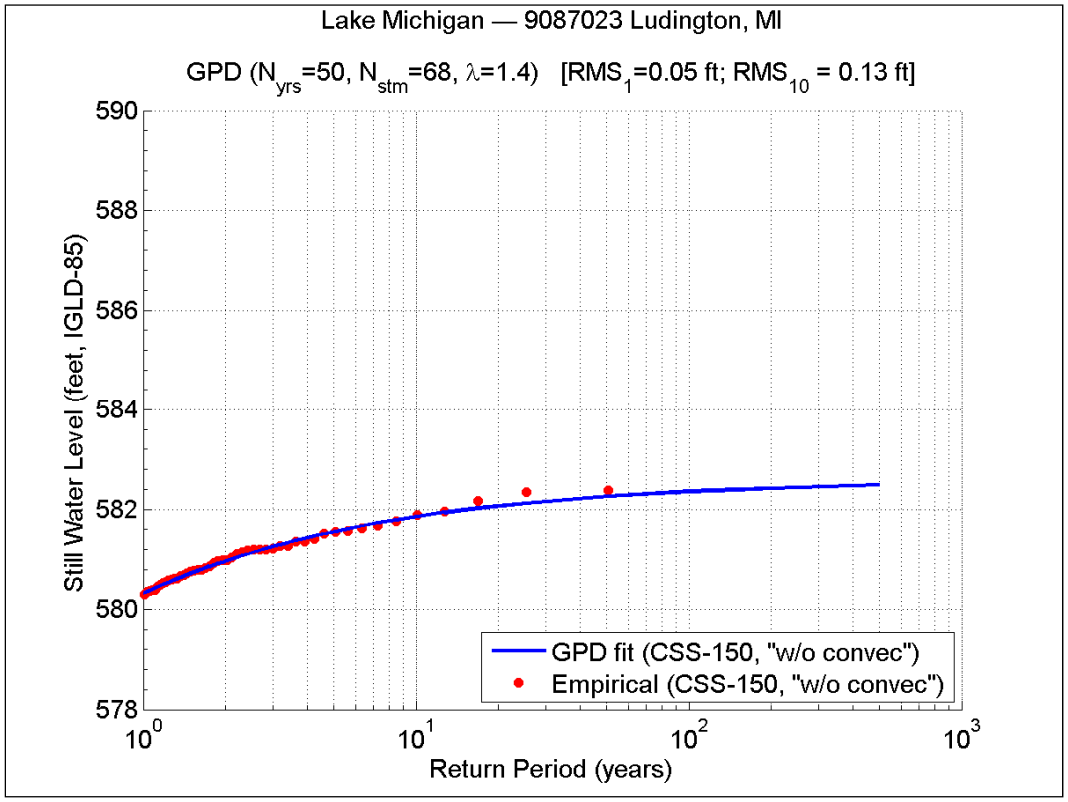



Figure 53. SWL RP plot from CSS (no convective storms) for Ludington, MI.

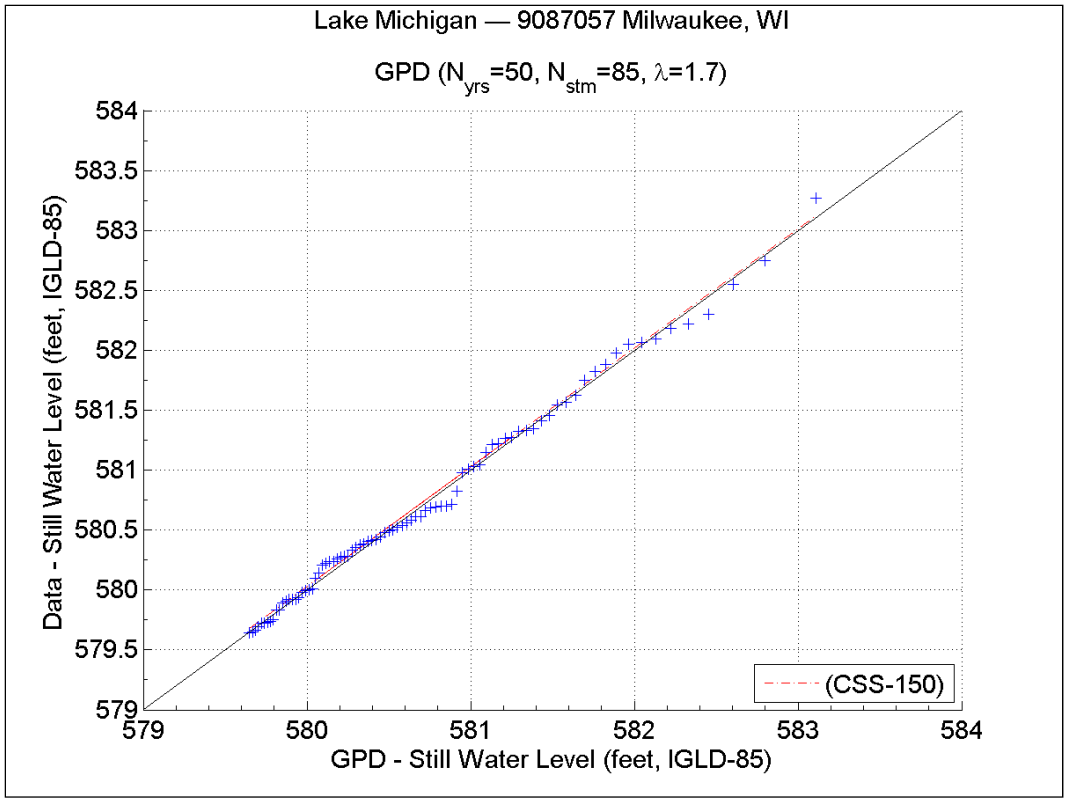



Figure 54. SWL Q-Q plot from CSS (with convective storms) for Milwaukee, WI.

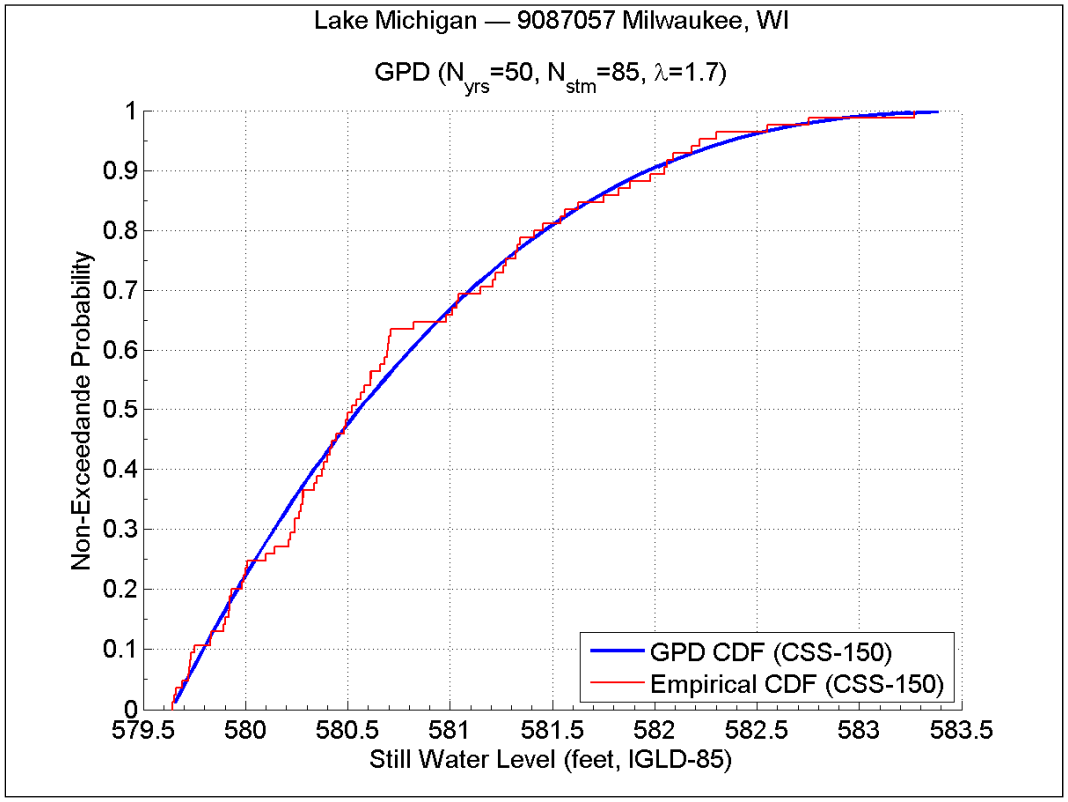



Figure 55. SWL CDF plot from CSS (with convective storms) for Milwaukee, WI.

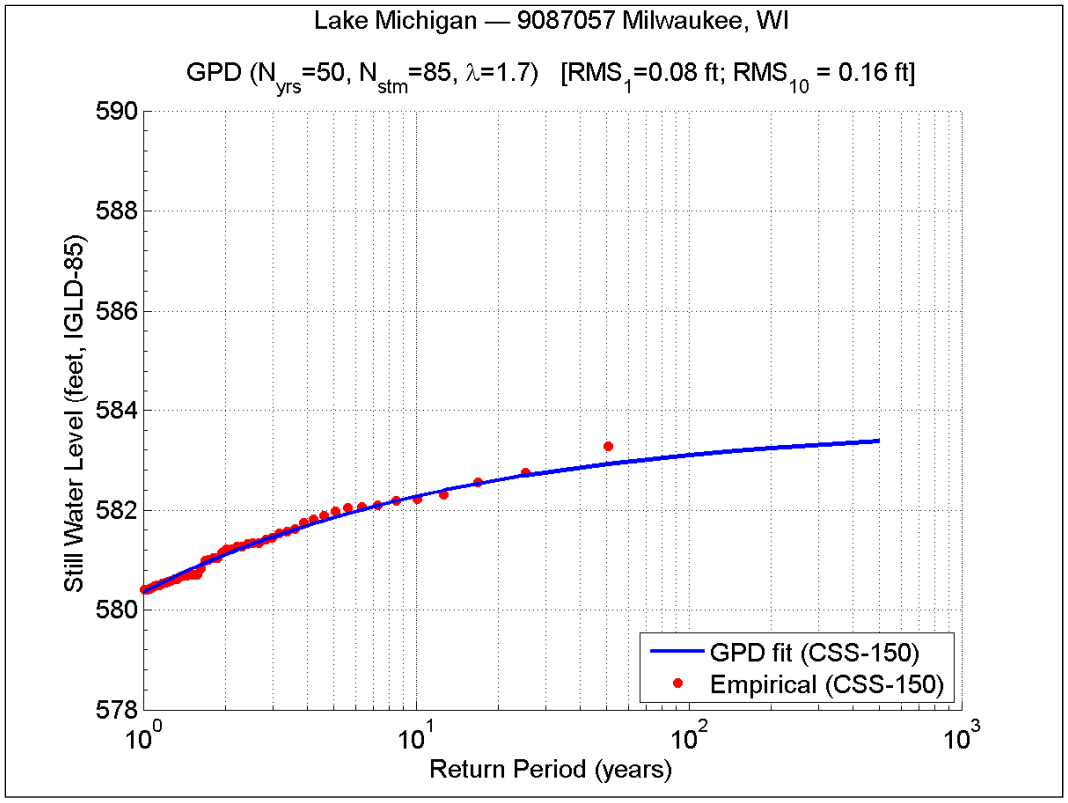



Figure 56. SWL RP plot from CSS (with convective storms) for Milwaukee, WI.

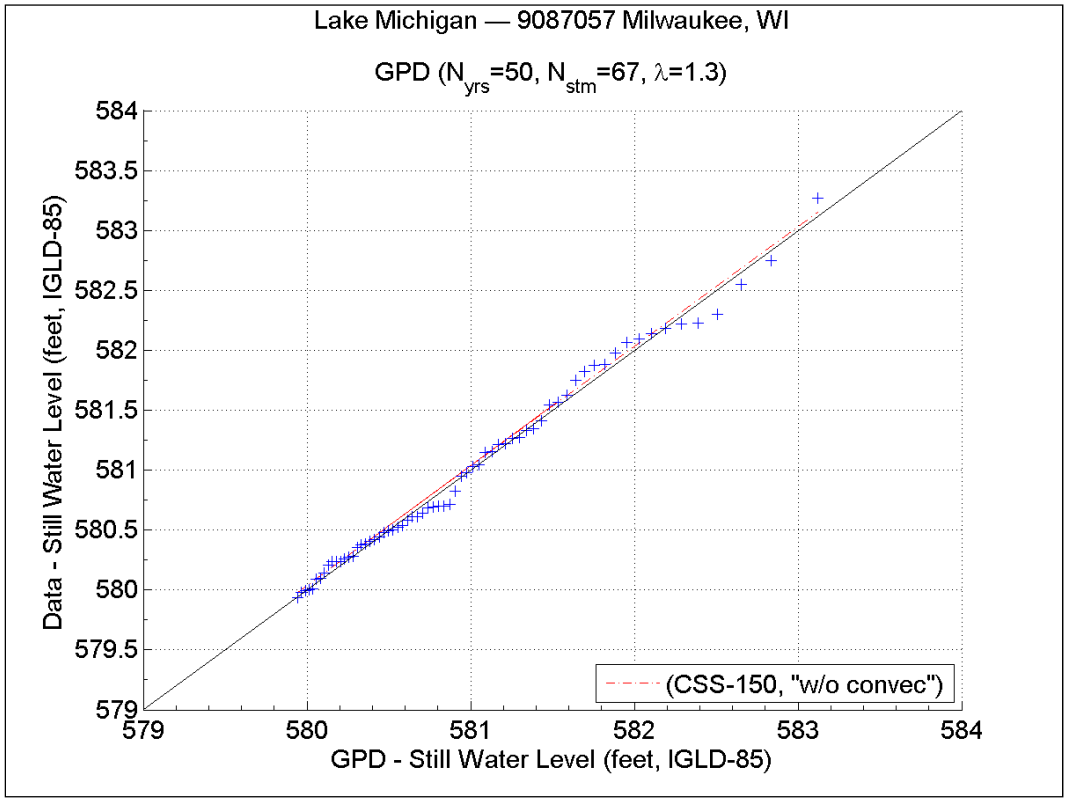



Figure 57. SWL Q-Q plot from CSS (no convective storms) for Milwaukee, WI.

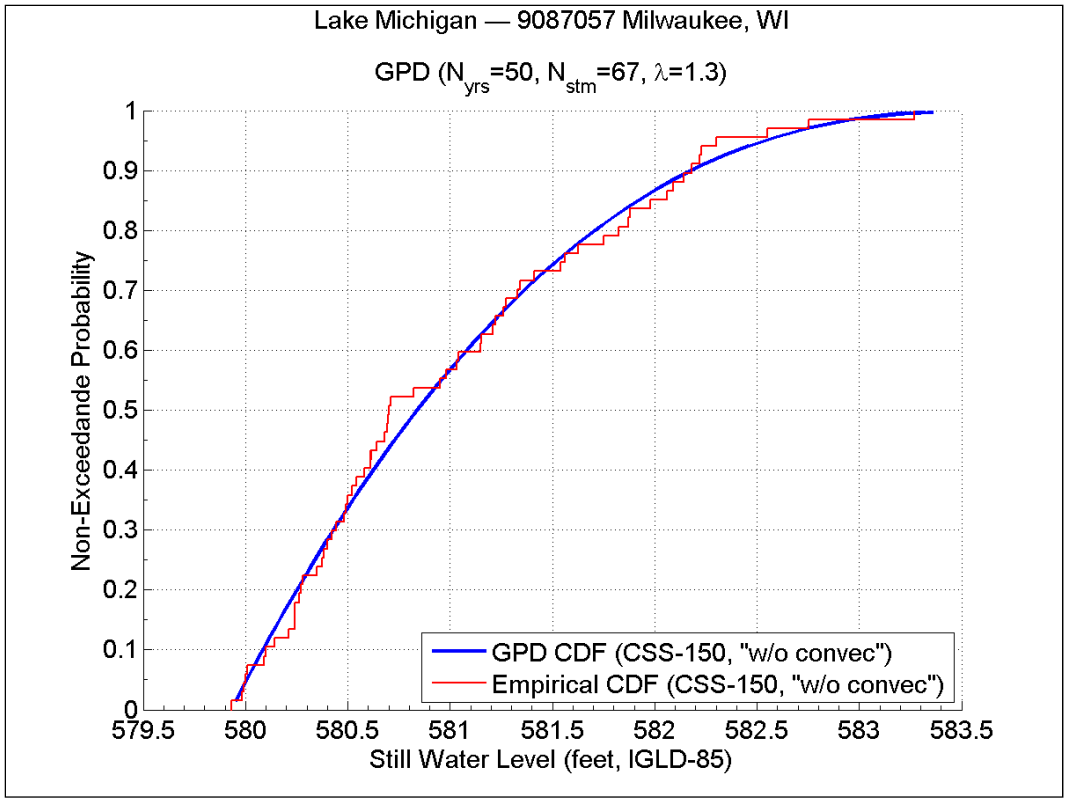



Figure 58. SWL CDF plot from CSS (no convective storms) for Milwaukee, WI.

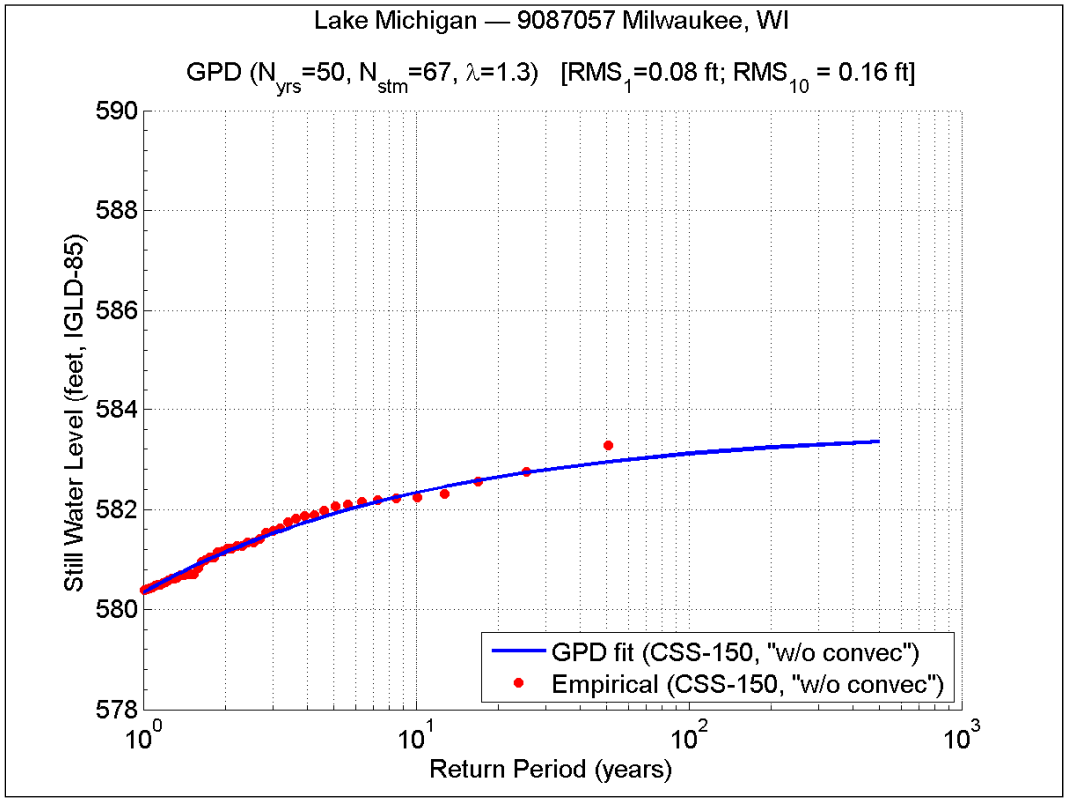



Figure 59. SWL RP plot from CSS (no convective storms) for Milwaukee, WI.

Table 6. SWL-CSS values for one percent and 0.2 percent annual chance events.

| SWL-CSS with summer events (ft, IGL-85) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Prob(%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 582.52 | 583.19 | 583.65 | 583.39 | 582.51 | 583.08 | 584.39 | 584.36 | 583.29 |
| 1 | 582.37 | 582.82 | 583.44 | 583.11 | 582.38 | 582.84 | 584.11 | 583.73 | 582.93 |
| SWL-CSS without summer events (ft, IGL-85) | | | | | | | | | |
| Prob(%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 582.49 | 583.11 | 583.54 | 583.36 | 582.37 | 582.29 | 584.34 | 584.26 | 583.21 |
| 1 | 582.36 | 582.85 | 583.35 | 583.12 | 582.27 | 582.22 | 584.1 | 583.71 | 582.91 |

Table 7. SWL-CSS vs. SWL-FSS values for one percent and 0.2 percent annual chance events.

| [(SWL-CSS with summer events) – (SWL-FSS with summer events)] (ft) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Prob(%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 0.12 | -0.10 | -0.02 | 0.15 | -0.69 | 0.06 | 0.22 | -0.88 | 0.12 |
| 1 | 0.05 | 0 | -0.02 | 0.04 | -0.55 | -0.03 | 0.06 | -0.45 | 0.05 |
| [(SWL-CSS without summer events) – (SWL-FSS with summer events)] (ft) | | | | | | | | | |
| Prob(%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 0.09 | -0.18 | -0.13 | 0.12 | -0.83 | -0.73 | 0.17 | -0.98 | 0.04 |
| 1 | 0.04 | 0.03 | -0.11 | 0.05 | -0.66 | -0.65 | 0.05 | -0.47 | 0.03 |

Regarding the third station, 9087072 (Sturgeon Bay, WI), the water level gage is located within the constricted section of the Surgeon Bay Canal which can influence the magnitude of the surge elevations and, thus, the SWL. The water level record length for Sturgeon Bay is just 60 years. Furthermore, Sturgeon Bay is also unusual in the sense that four of the top 10 surge events were convective storms. The Sturgeon Bay Canal might be more sensitive to convective events than other areas. To compensate for this inadequacy related to summer surge events, several wave-only storms for this area of Lake Michigan were added to the final 150-storm CSS, although waves produced by these events will not effectively penetrate far into Sturgeon Bay Canal. Aside from the Sturgeon Bay no significant differences were found after dropping convective events from the storm sets.

# 12 Total Water Level – Full Storm Set

The total water level (TWL) is equivalent to the still water level (SWL) with the addition of wave runup (TWL = SWL + runup). As part of this focus study, TWL distributions and return period plots were revised to incorporate updated runup estimates. It was found that Mase's equation (1989), used in preliminary calculations, often significantly overestimated runup magnitudes. Melby et al. (2012) recommended using either the modified Mase (1989) equation or the Stockdon equation (Stockdon et al. 2006) for runup on beaches. Herein, the Stockdon equation is used.

Runup and, hence, TWL computation for storm events occurring during the 1960s was based on monthly mean and maxima water levels and hourly surrogate wave estimates. All extreme surge events derived from 1960s water level monthly data were matched with the peak wave event occurring during the same month. In this way, 1960s storm surge events were assigned a specific date tied to synoptic extreme wave events.

One of the main objectives of this focus study was to determine the optimal prioritization ratio between water level events and wave events to be used for the storm sampling methodology. As part of the work by Melby et al. (2012) it was determined that a 50 percent storm surge to 50 percent wave height prioritization ratio resulted in accurate representation of the TWL distributions for the 27-year period (1970-1997) evaluated as part of that study. However, in light of internal technical team discussions regarding the storm selection approach, the validity of this premise was reevaluated for the full 50-year period from 1960 to 2010. In the current study, three prioritization ratios were evaluated: (1) 100 percent storm surge/0 percent wave runup; (2) 50 percent storm surge/50 percent wave runup; and (3) screening and selection on superimposed storm surge and wave runup (i.e., storm surge + wave runup).

In addition to the prioritization ratios, the significance of convective storms (summer events) was assessed in terms of how they influence the TWL distributions. For this purpose, all storm prioritization ratios were evaluated with and without the presence of convective storms. In the latter case, summer events were dropped from the FSS and replaced with non-convective storms.

## 12.1 Wave runup estimation

The Stockdon equation (Stockdon et al. 2006) was used to estimate wave runup elevations on beach profiles based on deepwater wave parameters, as follows:

$$R_{2\%} = 0.043(H_0 L_0)^{1/2} \tag{27}$$

for $\xi_0 < 0.3$

or,

$$R_{2\%} = 1.1\left(0.35\beta_f (H_0 L_0)^{1/2} + \frac{\left[H_0 L_0 \left(0.563\beta_f^2 + 0.004\right)\right]^{1/2}}{2}\right) \tag{28}$$

for $\xi_0 \geq 0.3$

where

$R_{2\%}$ 2% = exceedance value of wave runup (m, ft)
$\xi_0$ = offshore Iribarren number (non-dimensional)
$\beta_f$ = foreshore beach slope near SWL (non-dimensional)
$H_0$ = deepwater zero moment significant wave height (m, ft)
$L_0$ = deepwater wavelength (m, ft)

## 12.2 Surge/wave event prioritization

A preliminary assessment of the surge/wave prioritization ratio was performed early in the study. For this assessment, 150 top surge storms and 150 top wave storms were sampled, for a total of 300 storms, and different ratios were evaluated. It was found that the optimal surge/wave ratio can vary depending on the slope of the beach/structure being analyzed:

- For 1:80 to 1:30 slopes – optimal ratios varied slightly from 50 percent surge/ 50 percent waves to 40 percent surge/60 percent waves.
- For 1:5 slopes or steeper – optimal ratio is roughly 30 percent surge/70 percent waves, mainly because wave events tend to be more influential on runup for steep slopes.

For the majority of the cases, where the coast is characterized by natural gentle beach slopes, a 50 percent surge/50 percent wave ratio is adequate. However, in shallow coastal lowland areas with a wide flood plain, inundation and overland wave propagation, waves may have little influence on the extent of flood inundation. Waves become depth-limited as they propagate inland and are dampened by obstructions such as vegetation and buildings. By the time the wave reaches the point of intersection between the still water and ground, the significance of wave events is typically negligible. For this case a ratio closer to 100 percent surge/0 percent waves might be a better choice.

The critical issue is to assess how the different prioritization ratios actually affect the TWL extremal distributions and the estimation of one percent and 0.2 percent annual chance values. For the case of surge + runup, the Equations 27-28 were used to estimate wave runup elevations at every 1-hour time step for the entire 1960-2010 period, thus generating a complete time series of runup. However, the water level record length at some locations was less than 50 years (i.e., Kewaunee with 37 years, and Port Inland with 45 years). Since storm surge values are derived from measured water levels, the computation of runup and TWL at some locations was limited by the water level record length. After the runup and surge time series were superimposed, POT was performed to sample the extreme storms from this set. The TWL-FSS generated from the surge + runup time series are considered to represent the true distributions for comparison and analysis purposes.

TWL-FSS return period plots based on storm surge + wave runup (s+r) data are shown in Figures 60 and 61, for the Ludington gage. Note that Figure 60 included convective storms while Figure 61 excludes them. Similar plots for the Milwaukee gage are displayed in Figures 66-67. Return period plots corresponding to the 50 percent storm surge/50 percent wave runup ratio are illustrated in Figures 62-63, for Ludington; and Figures 68-69 for Milwaukee. Likewise, plots corresponding to 100 percent storm surge/ 0 percent wave runup ratio are displayed in Figures 64-65, for Ludington; Figures 70-71 for Milwaukee.

Note that the Q-Q plots and CDF plots associated with the TWL-FSS from Ludington and Milwaukee gages, are omitted from this section; but they are presented in Appendix A.

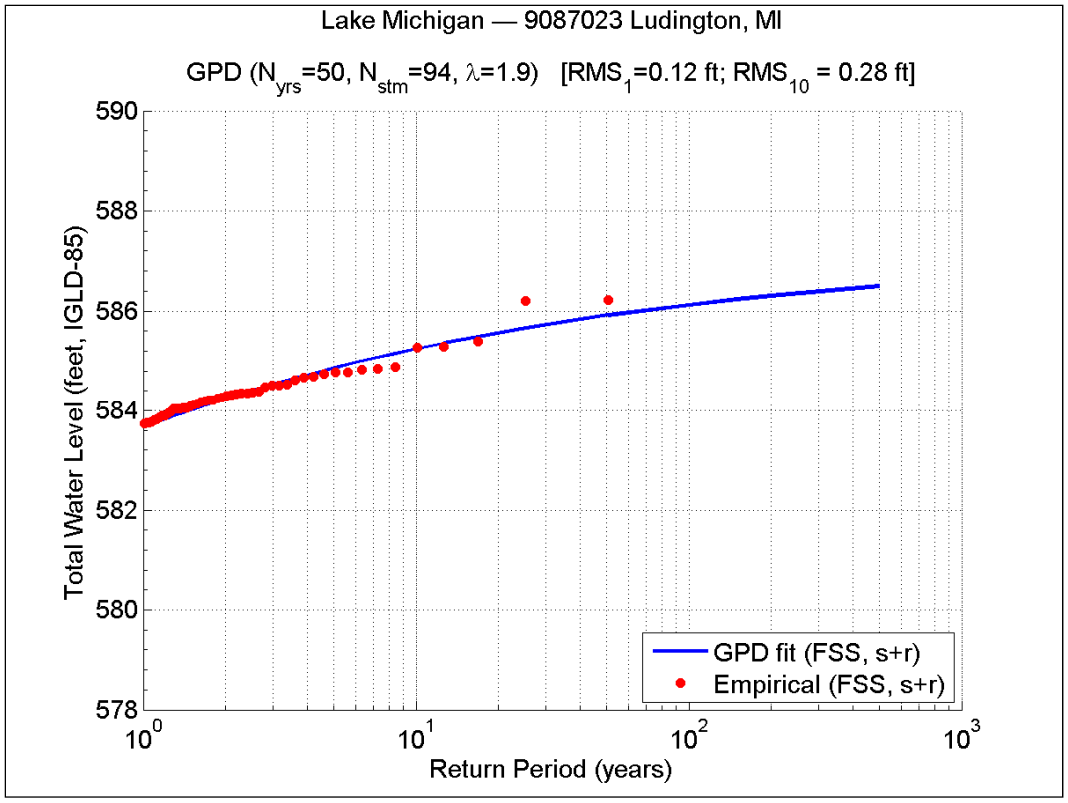



Figure 60. TWL RP plot from FSS (s+r, with convective storms) for Ludington, MI.

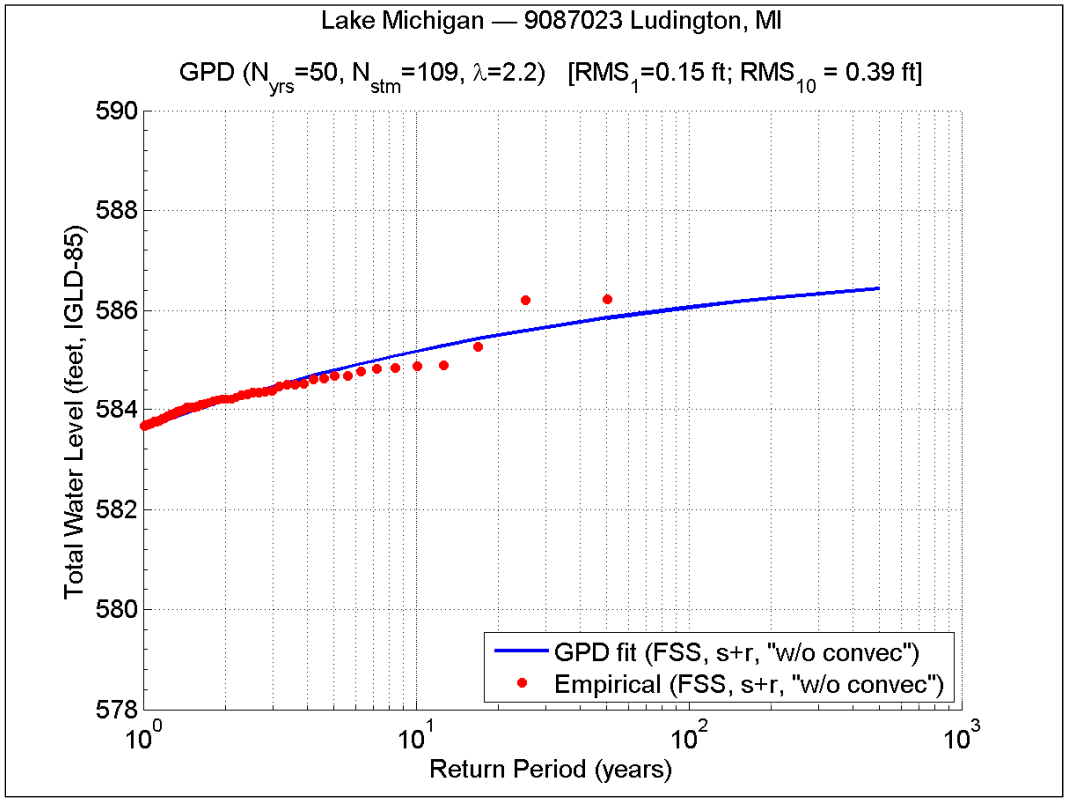



Figure 61. TWL RP plot from FSS (s+r, no convective storms) for Ludington, MI.

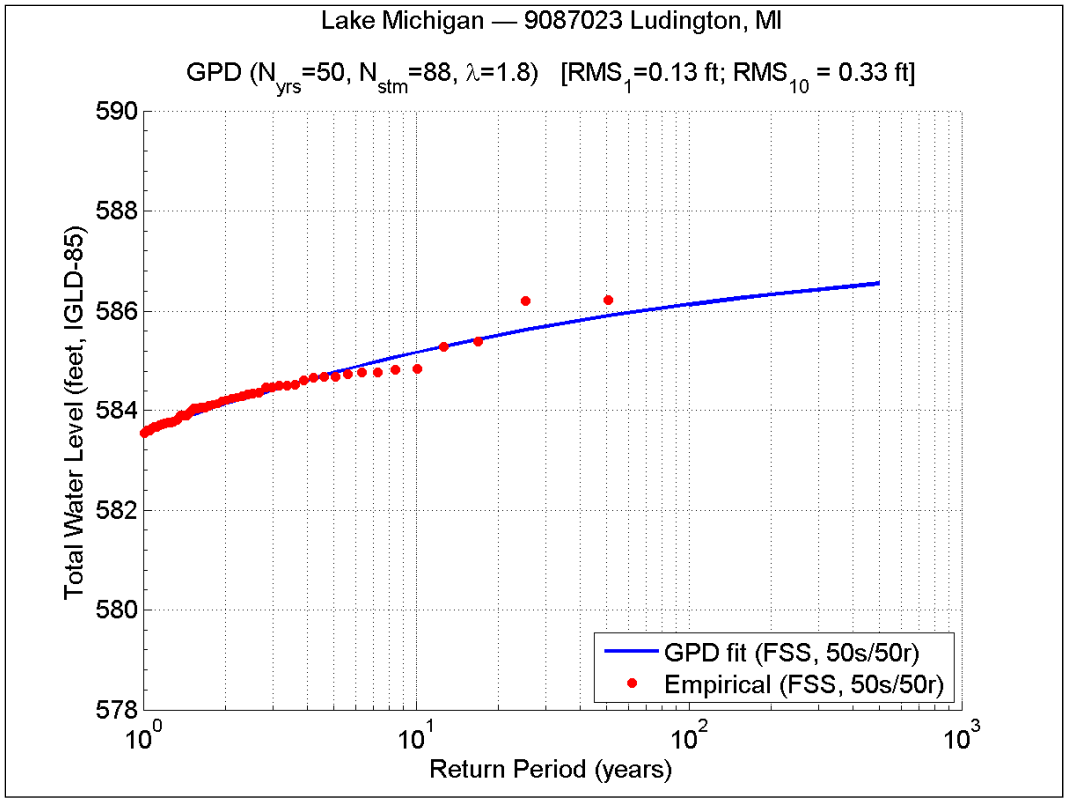



Figure 62. TWL RP plot from FSS (50s/50r, with convective storms) for Ludington, MI.

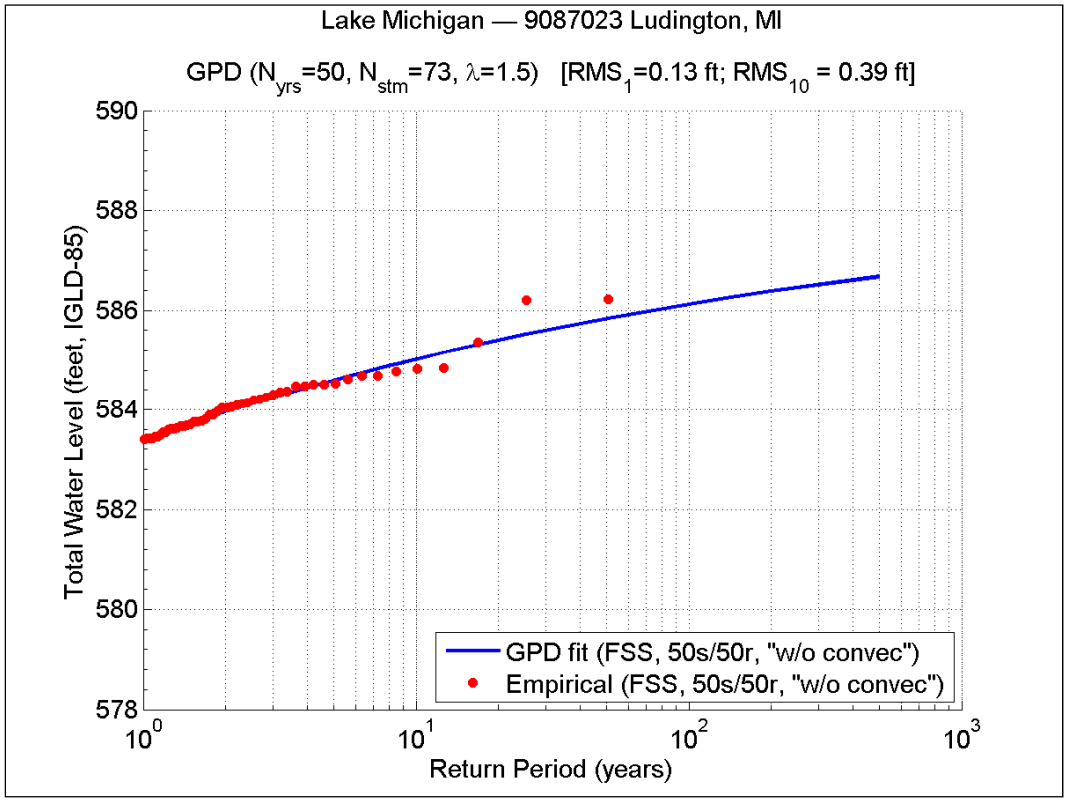



Figure 63. TWL RP plot from FSS (50s/50r, no convective storms) for Ludington, MI.

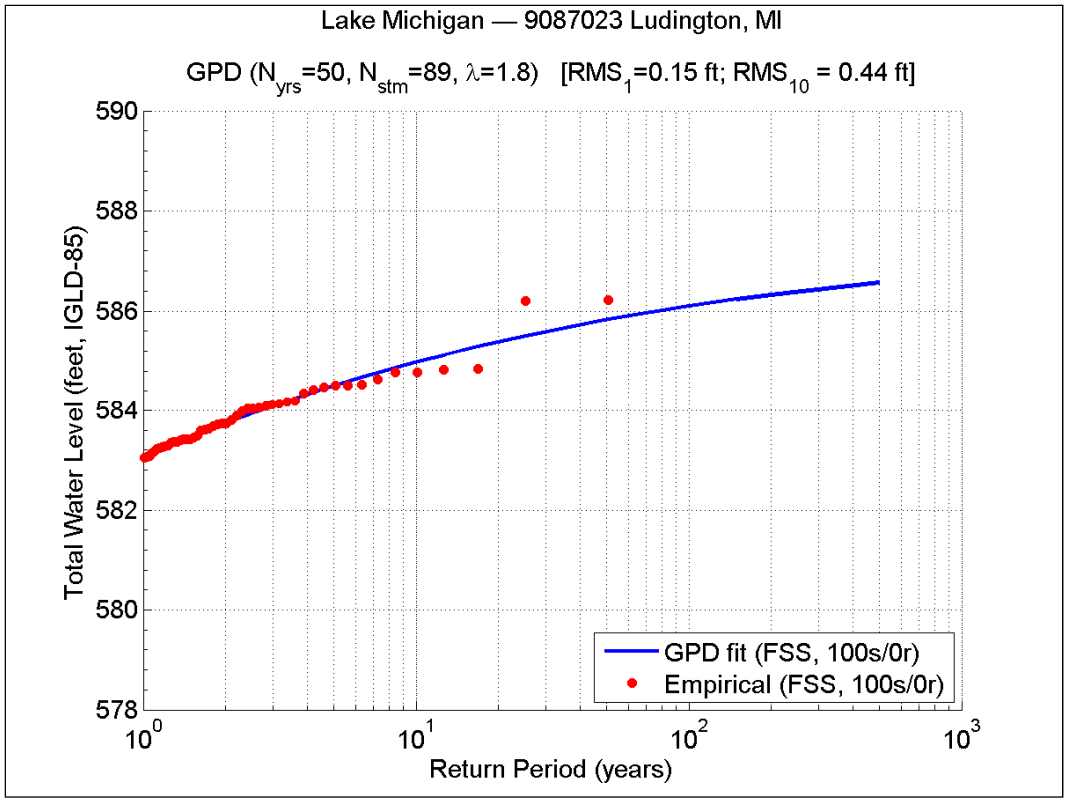



Figure 64. TWL RP plot from FSS (100s/0r, with convective storms) for Ludington, MI.

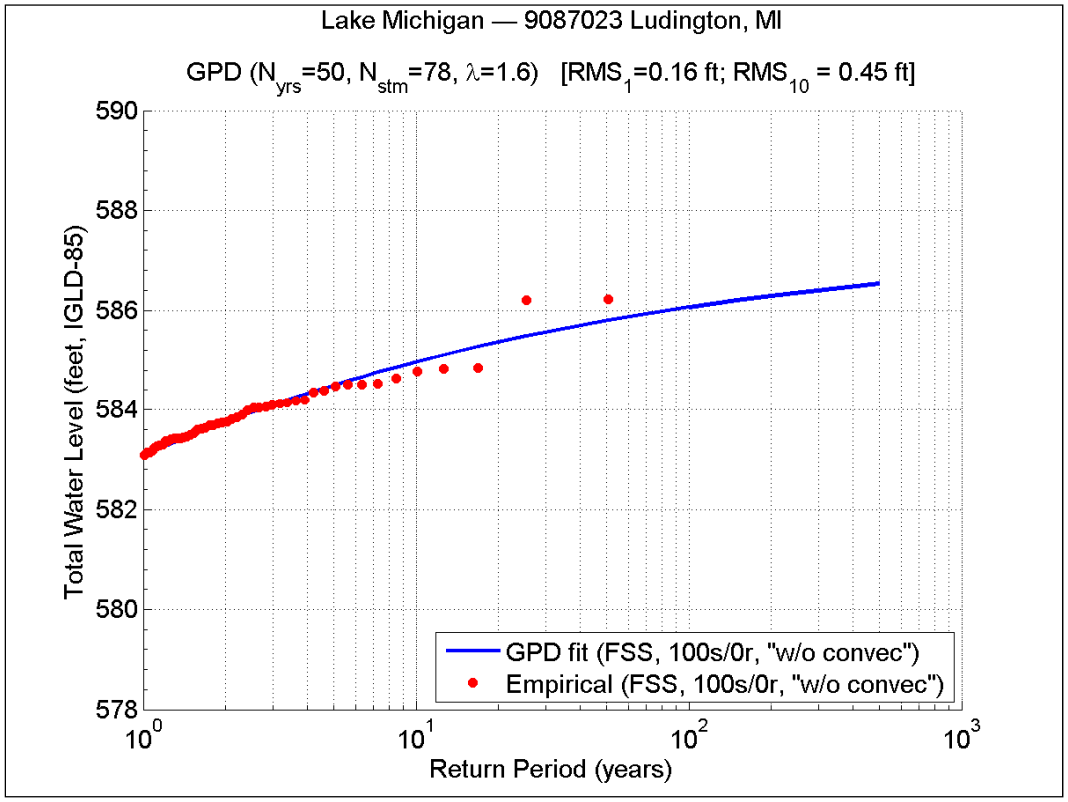



Figure 65. TWL RP plot from FSS (100s/0r, no convective storms) for Ludington, MI.

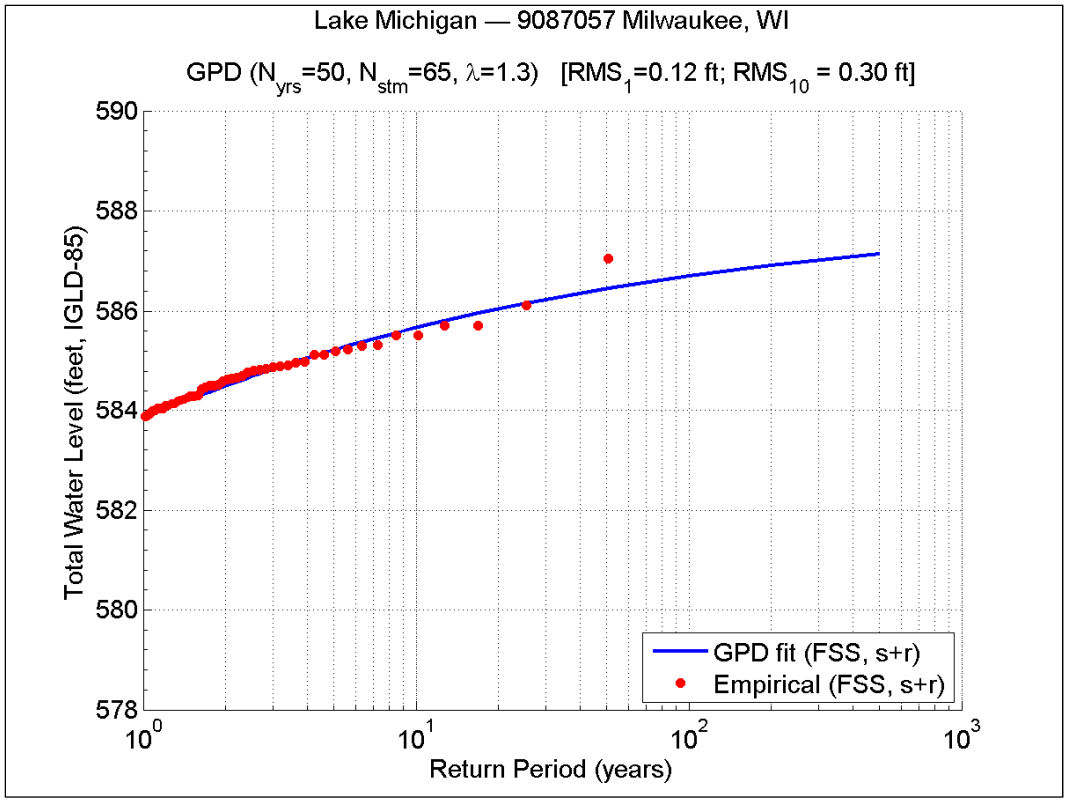



Figure 66. TWL RP plot from FSS (s+r, with convective storms) for Milwaukee, WI.

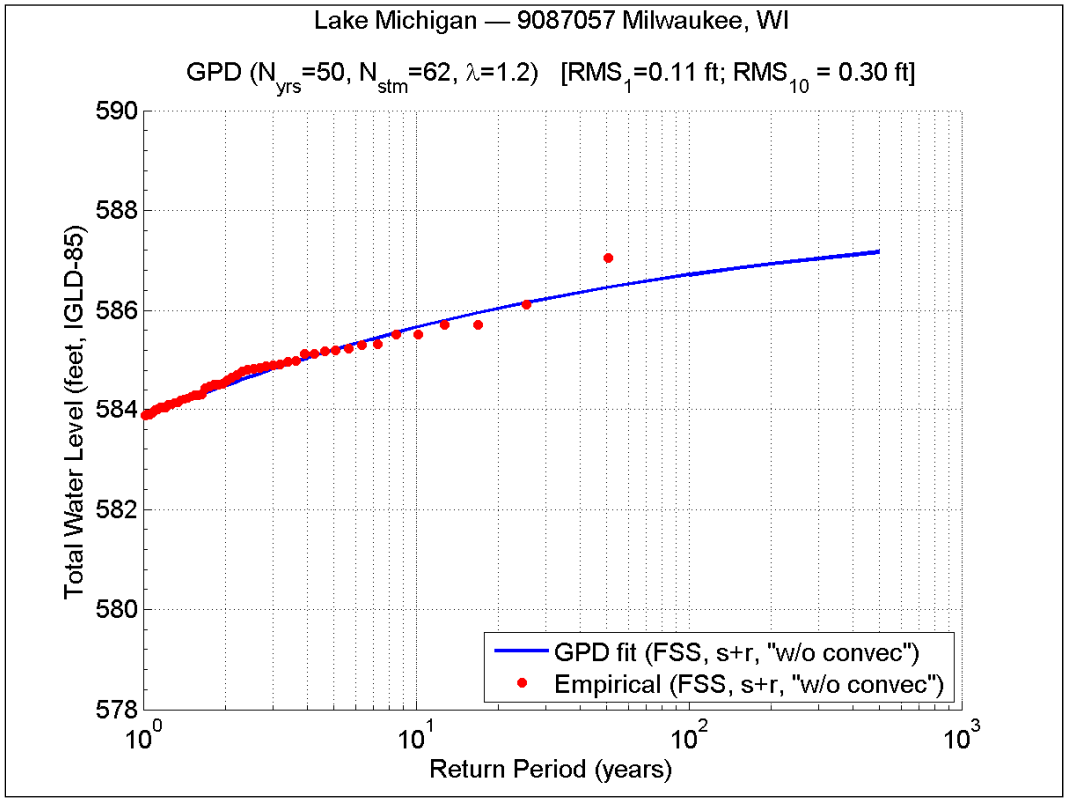



Figure 67. TWL RP plot from FSS (s+r, no convective storms) for Milwaukee, WI.

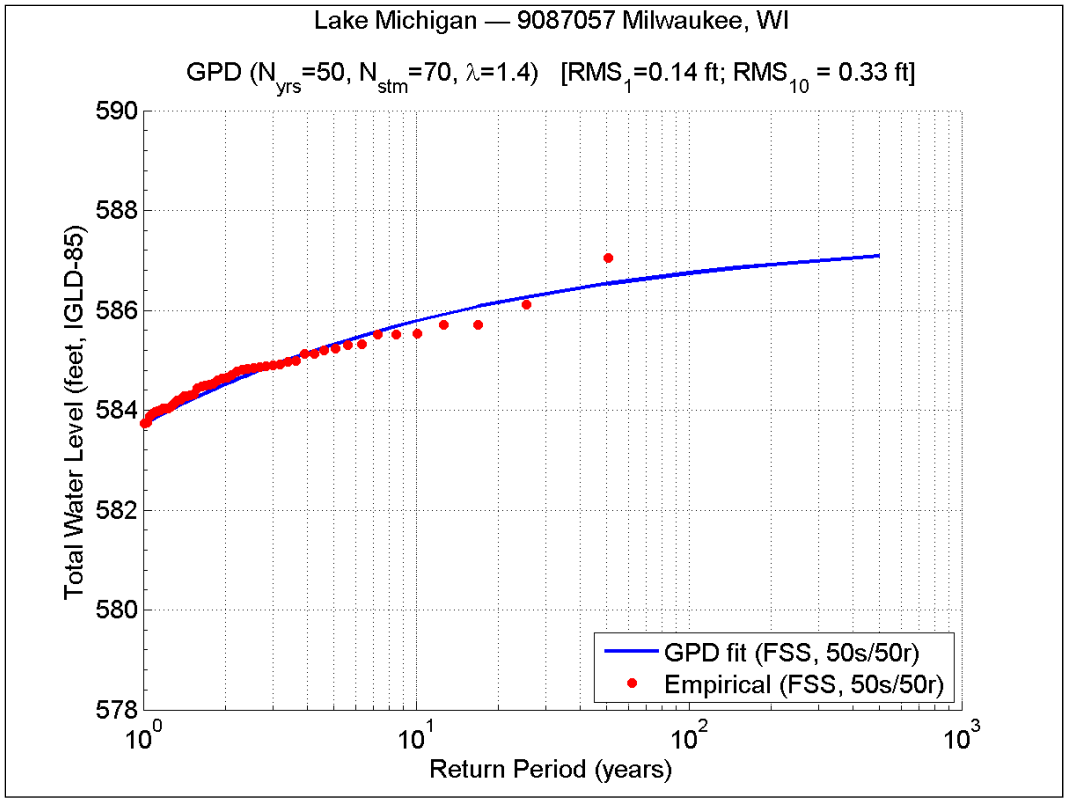



Figure 68. TWL RP plot from FSS (50s/50r, with convective storms) for Milwaukee, WI.

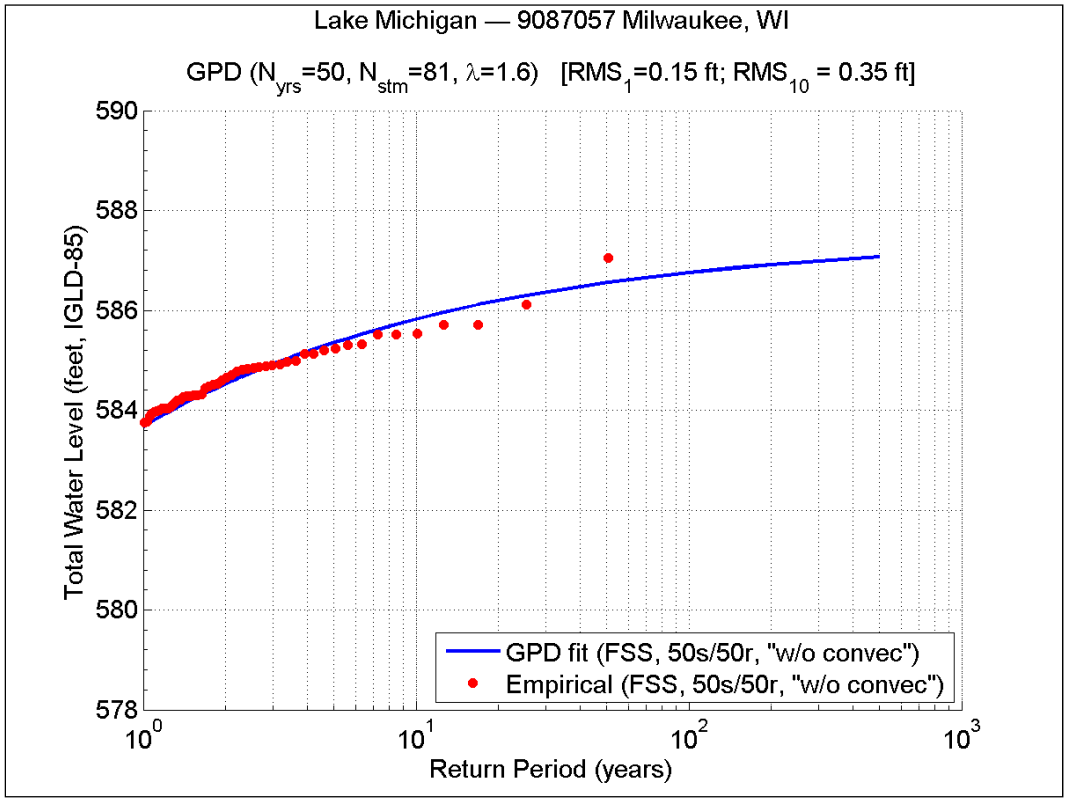



Figure 69. TWL RP plot from FSS (50s/50r, no convective storms) for Milwaukee, WI.

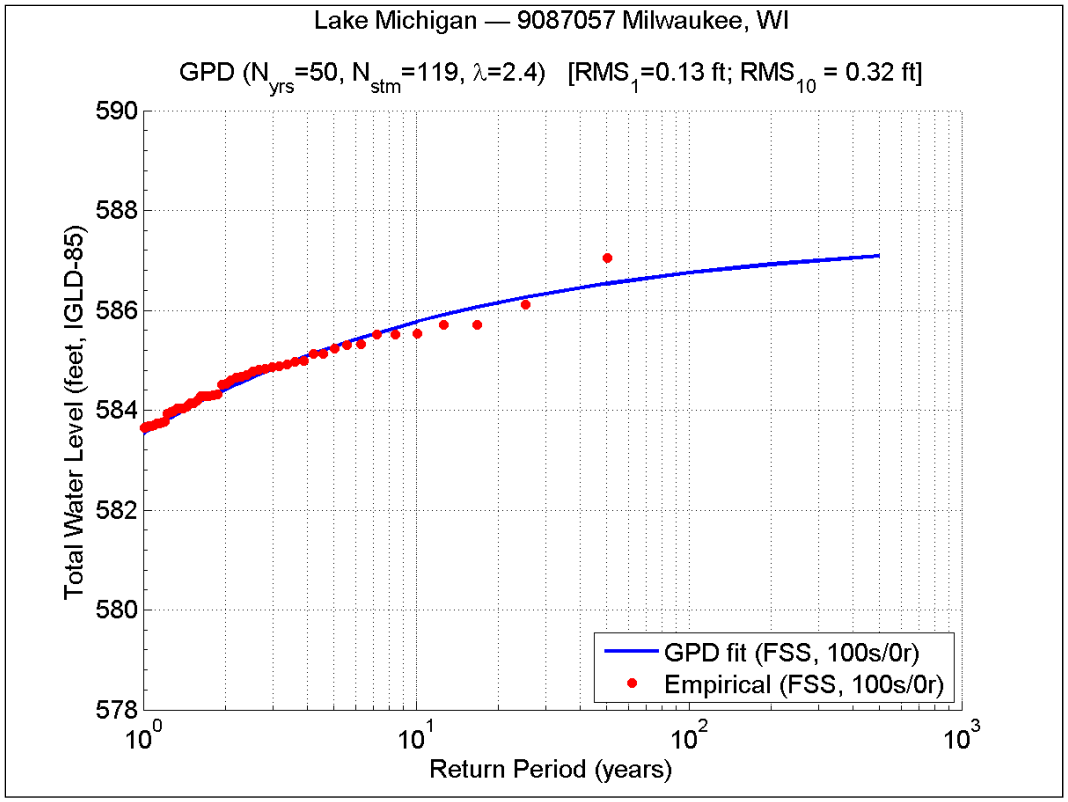



Figure 70. TWL RP plot from FSS (100s/0r, with convective storms) for Milwaukee, WI.

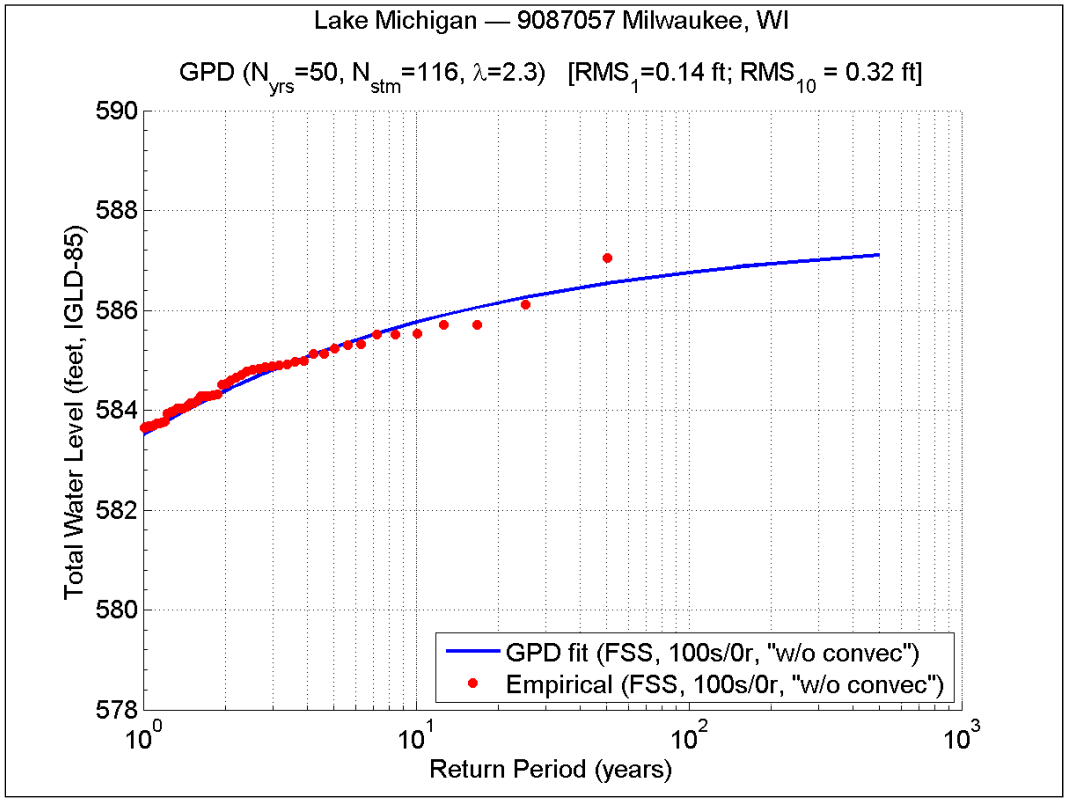



Figure 71. TWL RP plot from FSS (100s/0r, no convective storms) for Milwaukee, WI.

Table 8 shows the TWL-FSS estimates corresponding to one percent and 0.2 percent annual chance TWL values, both with and without convective storms, for all three of the different screening approaches:

Table 8. TWL-FSS values for on epercent and 0.2 percent annual chance events.

| TWL-FSS with summer events (surge + wave runup) (ft, IGLD-85) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| P(%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.51 | 585.99 | 586.98 | 587.15 | 585.59 | 585.42 | 587.17 | 588.88 | 585.06 |
| 1 | 586.13 | 585.82 | 586.88 | 586.71 | 585.37 | 585.22 | 586.82 | 587.95 | 584.58 |
| TWL-FSS without summer events (surge + wave runup) (ft, IGLD-85) | | | | | | | | | |
| P(%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.45 | 585.98 | 586.95 | 587.18 | 585.95 | 585.48 | 587.22 | 588.85 | 584.87 |
| 1 | 586.07 | 585.77 | 586.88 | 586.72 | 585.51 | 585.14 | 586.82 | 587.88 | 584.56 |
| TWL-FSS with summer events (50% surge - 50% wave) (ft, IGLD-85) | | | | | | | | | |
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.58 | 585.78 | 586.99 | 587.10 | 585.41 | 585.43 | 587.20 | 589.09 | 585.44 |
| 1 | 586.10 | 585.58 | 586.88 | 586.76 | 585.19 | 585.11 | 586.84 | 588.06 | 584.67 |
| TWL-FSS without summer events (50% surge - 50% wave) (ft, IGLD-85) | | | | | | | | | |
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.54 | 585.93 | 586.98 | 587.11 | 585.41 | 585.44 | 587.45 | 589.57 | 585.61 |
| 1 | 586.07 | 585.69 | 586.87 | 586.77 | 585.19 | 585.10 | 586.94 | 588.19 | 584.70 |
| TWL-FSS with summer events (100% surge - 0% wave) (ft, IGLD-85) | | | | | | | | | |
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.56 | 585.92 | 586.95 | 587.10 | 585.19 | 585.39 | 587.14 | 589.36 | 585.03 |
| 1 | 586.14 | 585.73 | 586.88 | 586.75 | 585.09 | 585.24 | 586.83 | 588.13 | 584.42 |
| TWL-FSS without summer events (100% surge - 0% wave) (ft, IGLD-85) | | | | | | | | | |
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.68 | 585.90 | 586.95 | 587.08 | 585.11 | 585.38 | 587.25 | 589.31 | 585.12 |
| 1 | 586.13 | 585.73 | 586.88 | 586.76 | 585.00 | 585.16 | 586.87 | 588.03 | 584.47 |

From these results, it can be concluded that the effect of rejecting and replacing convective storms is minimal on the TWL distributions. In the case of TWL-FSS corresponding to 50 percent surge/50 percent waves ratio, the average difference for all stations considered was 0.11 ft for 0.2 percent annual chance values and 0.03 ft for one percent annual chance values. On

average, comparisons between TWL-FSS with and without summer convective events showed differences of just 0.02 to 0.04 ft.

Comparisons between the 50 percent surge/50 percent waves and 100 percent surge/0 percent waves ratios and the surge + runup (close to true distribution) showed similarly negligible differences. In fact, the differences on TLW were typically around 0.03 ft, in average.

# 13 Total Water Level – Composite Storm Set

The TWL-CSS is defined as a single set, or composite, of storms derived from storm samples from different gage locations around a given lake. For the 50 year period, between 1960 and 2010, it was determined that the minimum number of total storms (per lake) that should be used to construct the CSS is 150, to achieve adequate overall distribution shapes and accurate estimates of TWL corresponding to one percent and 0.2 percent annual chance values. Similar to the previous analysis of storm water level, the CSS was also evaluated for total water level estimates. The TWL-CSS was generated following the same methodology that was used to generate the SWL-CSS, as discussed in Section 11; the only difference being that TWL data includes wave runup superimposed on SWL.

TWL-CSS return period plots corresponding to 50 percent storm surge/ 50 percent wave runup ratio, with and without summer convective storms, are illustrated in Figures 72-73, for Ludington; and Figures 76-77 for Milwaukee. Likewise, plots corresponding to 100 percent storm surge/ 0 percent wave runup ratio are displayed in Figures 74-75, for Ludington; and Figures 78-79 for Milwaukee.

Note that the Q-Q plots and CDF plots associated with the TWL-CSS from Ludington and Milwaukee gages, are omitted from this section; they are presented in Appendix B.

Table 9 shows the TWL-CSS estimates corresponding to one percent and 0.2 percent annual chance values, both with and without convective storms, for the two screening approaches:

Similarly to the TWS-FSS results, it can be concluded that the effect of rejecting and replacing convective storms in the case of TWL-CSS is minimal. On average, comparisons between TWL-CSS with and without convective events showed differences of just 0.04 to 0.17 ft.

Comparisons between the 50 percent surge/50 percent waves and 100 percent surge/0 percent waves ratios showed similarly negligible differences. The differences on TLW were typically between 0.05 ft and 0.15 ft, on average.

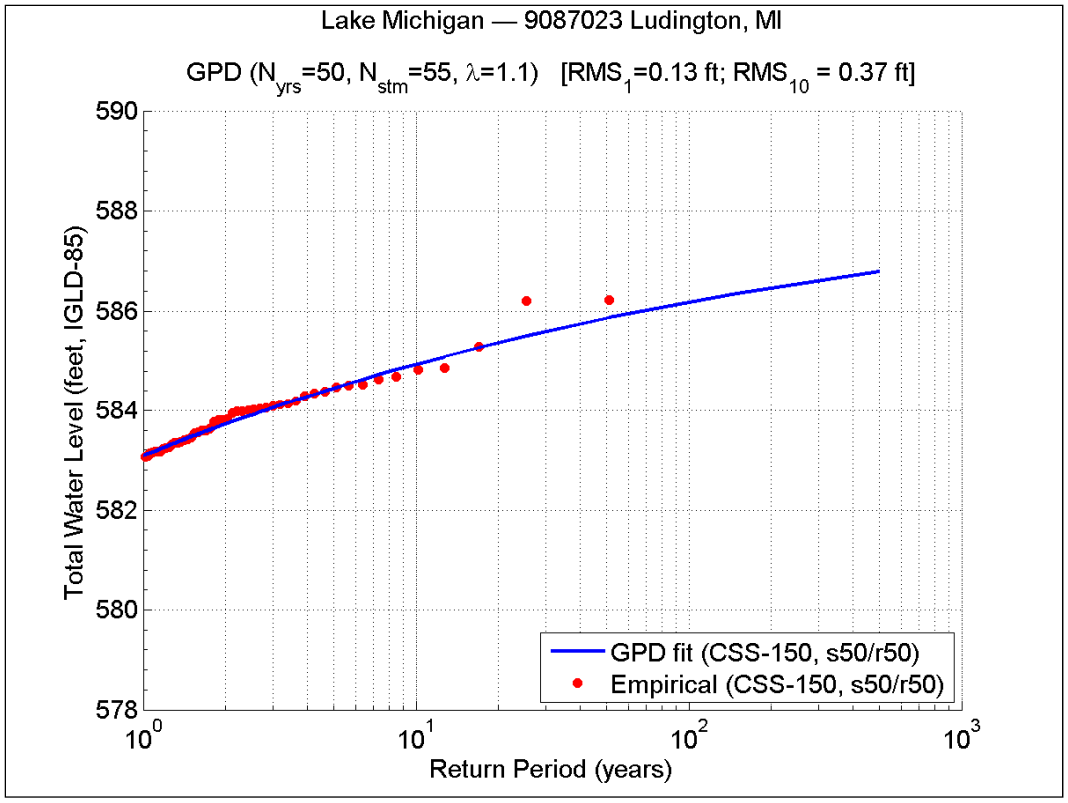



Figure 72. TWL RP plot from CSS (50s/50r, with convective storms) for Ludington, MI.

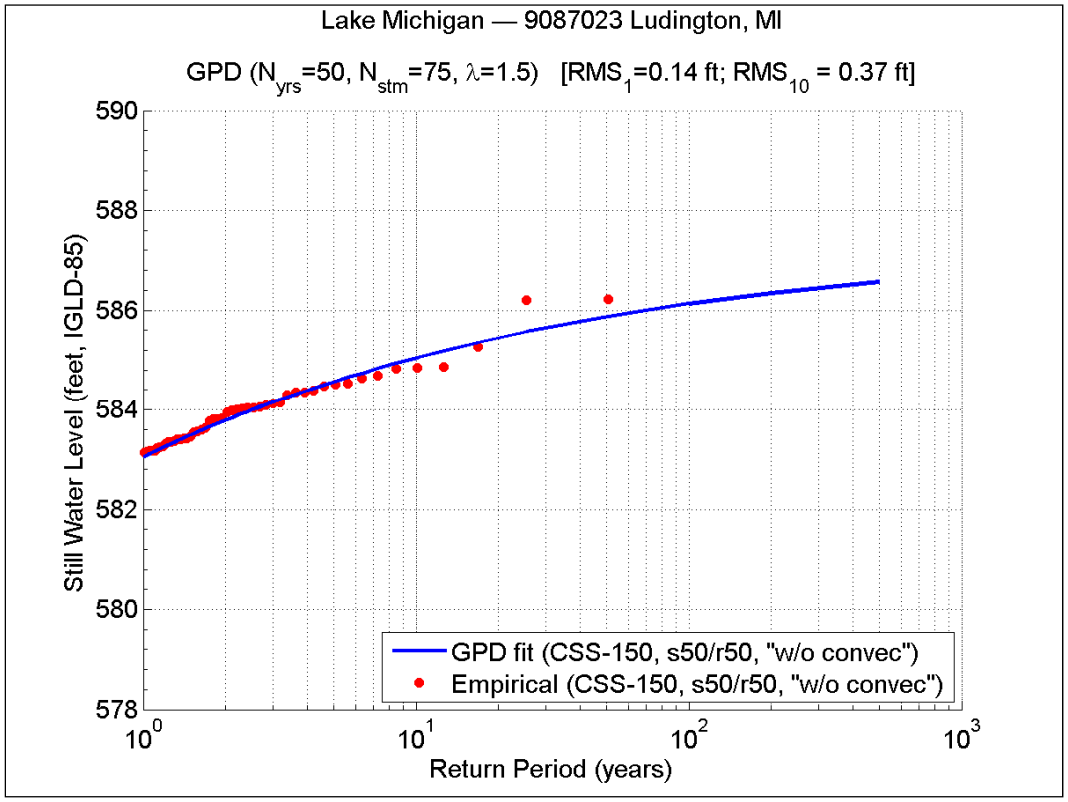



Figure 73. TWL RP plot from CSS (50s/50r, no convective storms) for Ludington, MI.

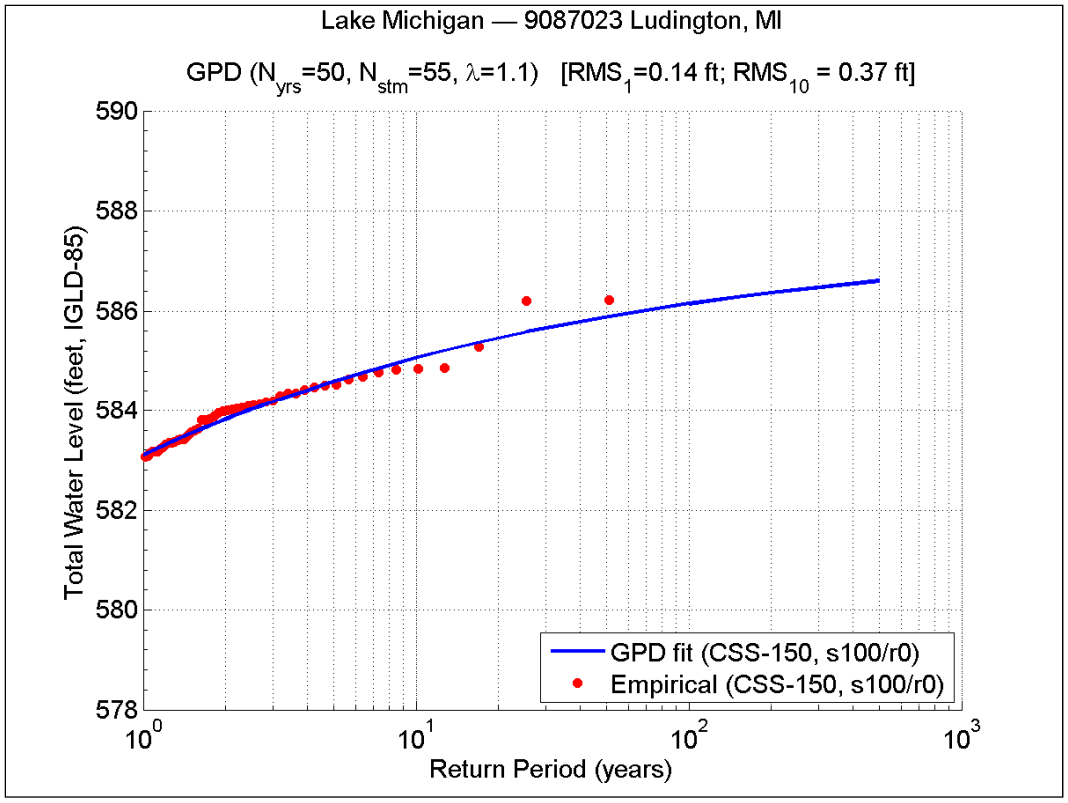



Figure 74. TWL RP plot from CSS (100s/0r, with convective storms) for Ludington, MI.

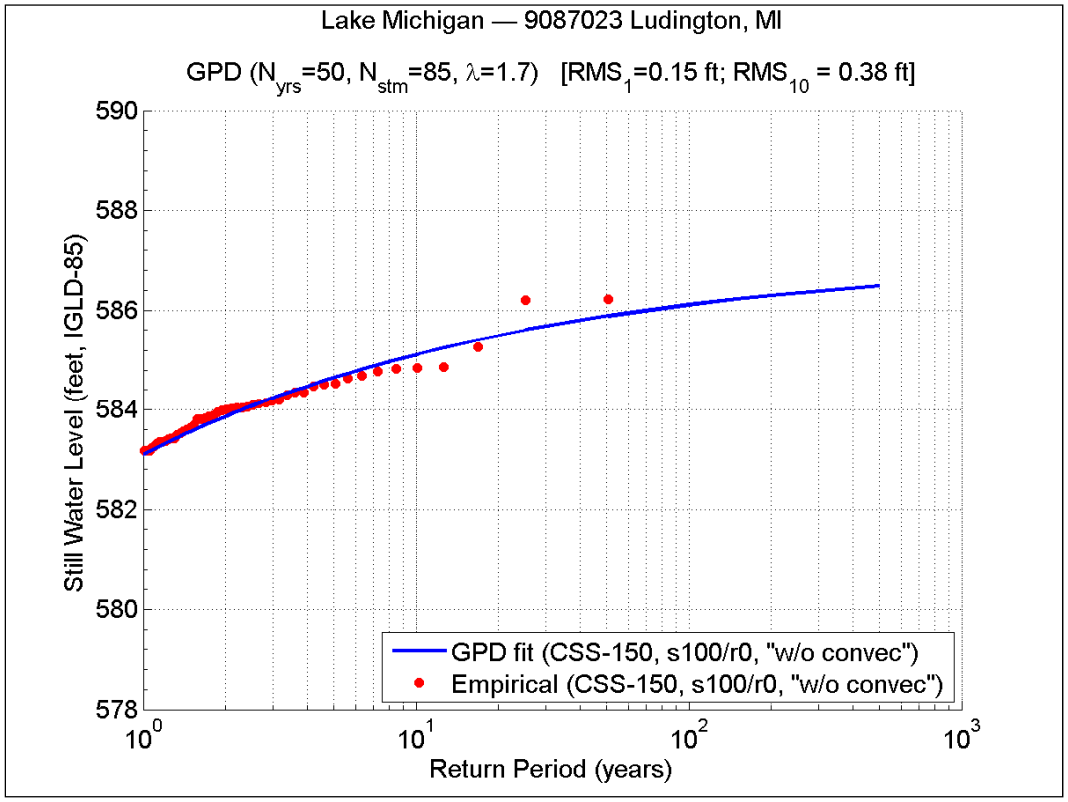



Figure 75. TWL RP plot from CSS (100s/0r, no convective storms) for Ludington, MI.

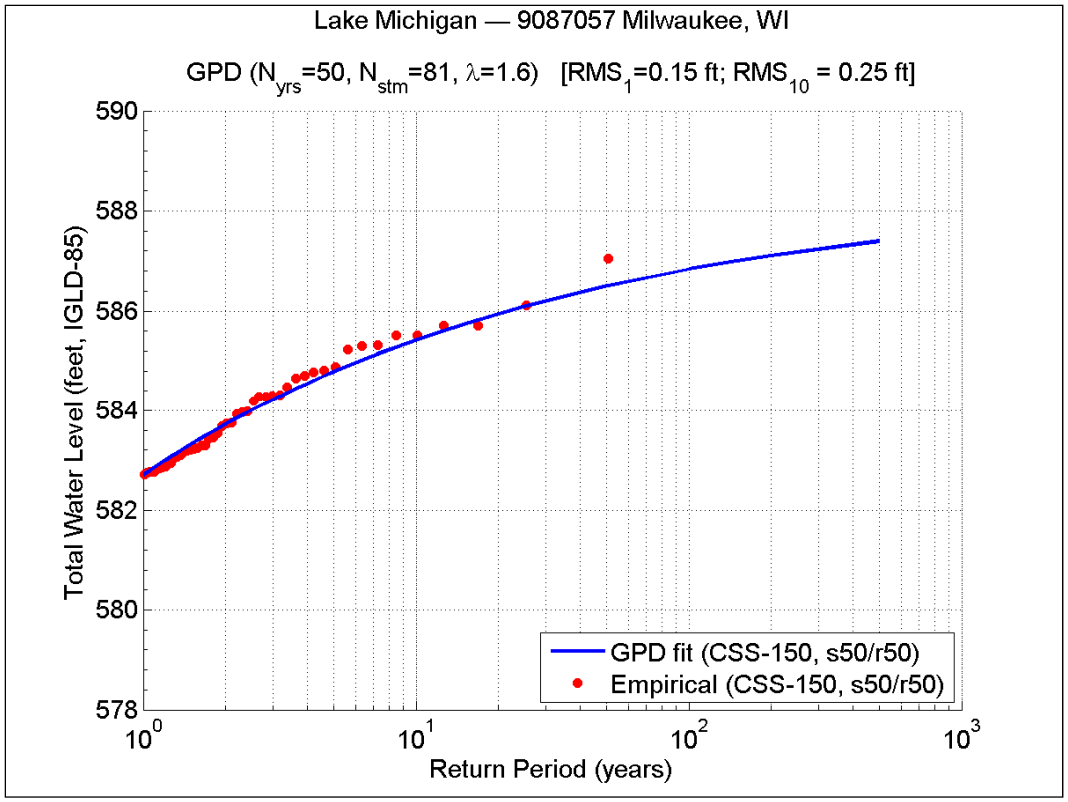



Figure 76. TWL RP plot from CSS (50s/50r, with convective storms) for Milwaukee, WI.

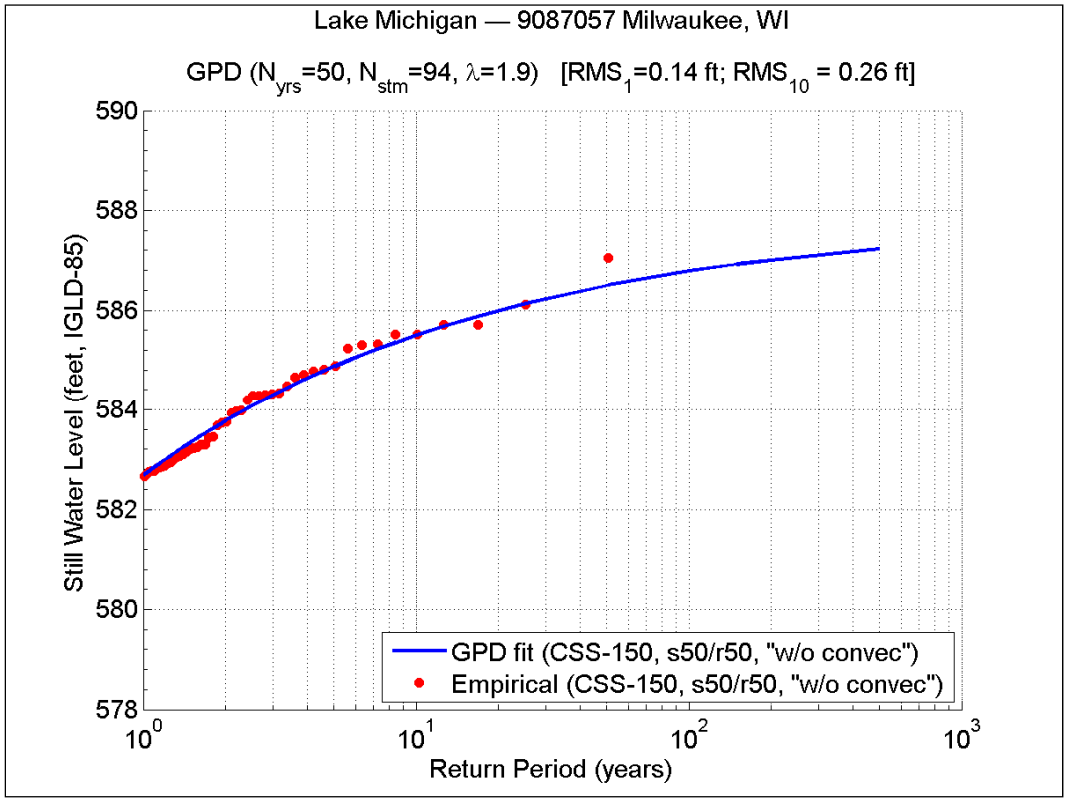



Figure 77. TWL RP plot from CSS (50s/50r, no convective storms) for Milwaukee, WI.

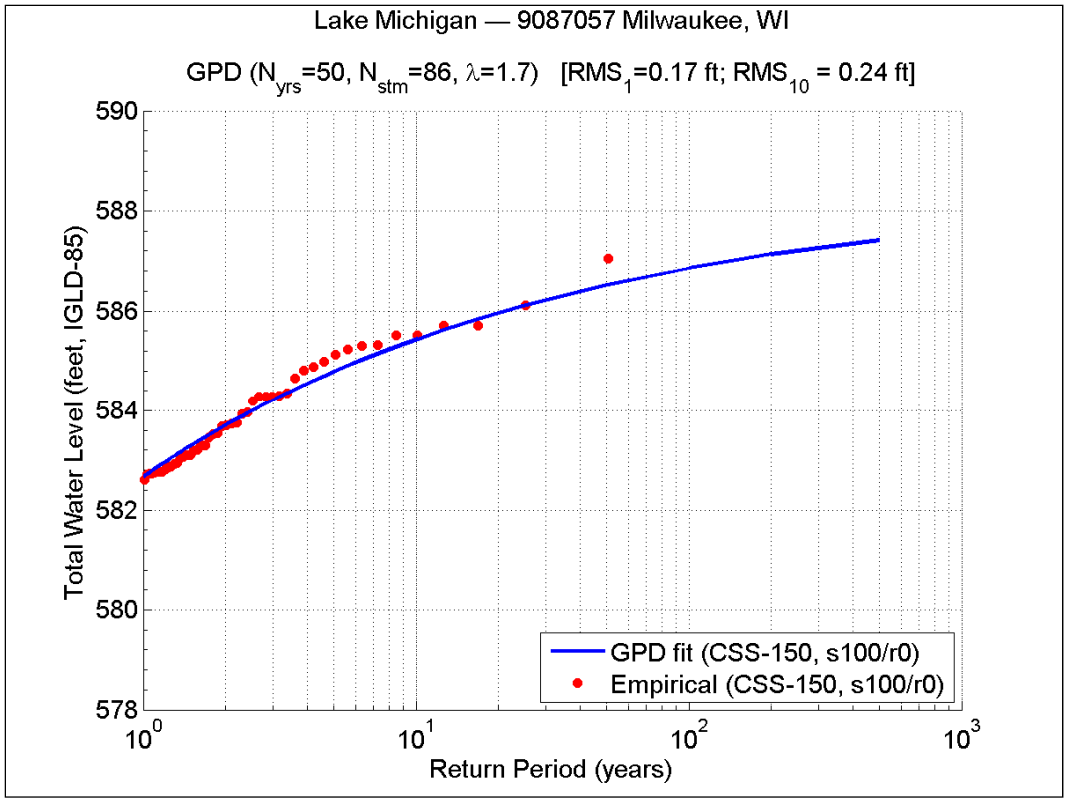



Figure 78. TWL RP plot from CSS (100s/0r, with convective storms) for Milwaukee, WI.

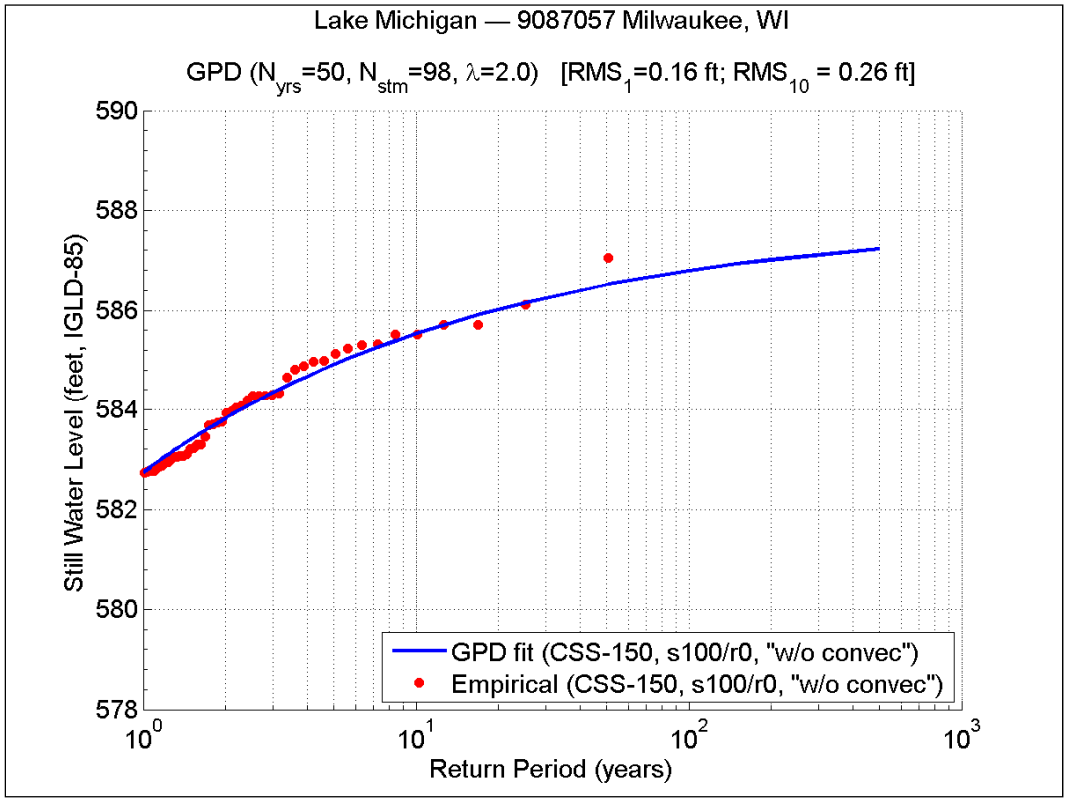



Figure 79. TWL RP plot from CSS (100s/0r, no convective storms) for Milwaukee, WI.

Table 9. TWL-CSS values for one percent and 0.2 percent annual chance events.

| TWL-CSS with summer events {50% surge - 50% wave} (ft, IGLD-85) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.79 | 586.07 | 587.1 | 587.4 | 585.22 | 585.38 | 587.7 | 588.8 | 585.13 |
| 1 | 586.18 | 585.73 | 586.87 | 586.84 | 585 | 585.18 | 587.07 | 587.99 | 584.49 |
| TWL-CSS without summer events {50% surge - 50% wave} (ft, IGLD-85) | | | | | | | | | |
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.58 | 585.98 | 587.1 | 587.24 | 585.14 | 585.37 | 587.2 | 588.76 | 584.88 |
| 1 | 586.13 | 585.69 | 586.88 | 586.79 | 584.96 | 585.15 | 586.83 | 587 | 584.39 |
| TWL-CSS with summer events {100% surge - 0% wave} (ft, IGLD-85) | | | | | | | | | |
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.61 | 586.01 | 586.74 | 587.43 | 585.21 | 585.37 | 587.28 | 588.81 | 584.95 |
| 1 | 586.16 | 585.72 | 586.58 | 586.86 | 585.04 | 585.22 | 586.89 | 588.01 | 584.42 |
| TWL-CSS with summer events {100% surge - 0% wave} (ft, IGLD-85) | | | | | | | | | |
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 586.5 | 585.92 | 586.72 | 587.24 | 585.11 | 585.38 | 587.24 | 588.78 | 584.82 |
| 1 | 586.12 | 585.68 | 586.61 | 586.8 | 584.99 | 585.14 | 586.88 | 587.978 | 584.38 |

## 13.1 Total water level: full storm set vs. composite storm set

The comparison between TWL-CSS and TWL-FSS is shown on Table 10. Table 10.a shows the differences between TWL-CSS with 50 percent surge/50 percent waves prioritization ratio without convective storms and TWL-FSS derived from surge plus runup POT with convective summer storms, for the one percent and 0.2 percent annual chance events. Table 10.b shows the differences between TWL-CSS with 100 percent surge/0 percent waves prioritization ratio without convective events and TWL-FSS derived from surge plus runup POT with convective storms, also for the one percent and 0.2 percent annual chance events. These TWL-CSS represent the results from the statistical analysis and storm sampling approach. TWL-FSS values with summer convective storms represent the true TWL distribution.

With the exception of station 9087068 (Kewaunee, WI), most differences in TWL are typically below 0.20 ft, with an average underprediction of 0.05 ft. The physical peculiarities of the Sturgeon Bay station were previously discussed in Section 11 (SWL-CSS). Also worth noting is the fact that no significant differences were found when comparing both the TWL-CSS with

50 percent surge/50 percent wave ratio and the TWL-CSS with 100 percent surge/0 percent wave ratio against the TWL-FSS (true distribution). The average difference between TWL-CSS and the true distribution with 50 percent surge/50 percent wave ratio, excluding station 9087068, was found to be -0.03 ft; likewise, average difference between TWL-CSS and the true distribution with 100 percent surge/0 percent wave ratio was estimated at -0.07 ft.

Table 10. TWL-CSS vs. TWL-FSS values for one percent and 0.2 percent annual chance events.

| [(TWL-CSS without summer events {50/50}) – (TWL-FSS with summer events {s+r})] (ft) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | 0.07 | -0.01 | 0.12 | 0.09 | -0.45 | -0.05 | 0.03 | -0.12 | -0.18 |
| 1 | 0 | -0.14 | 0 | 0.08 | -0.41 | -0.06 | 0.02 | -0.05 | -0.2 |
| [(TWL-CSS without summer events {100/0}) – (TWL-FSS with summer events {s+r})] (ft) | | | | | | | | | |
| P (%) | 87023 | 87031 | 87044 | 87057 | 87068 | 87072 | 87079 | 87096 | 75080 |
| 0.2 | -0.01 | -0.07 | -0.26 | 0.09 | -0.48 | -0.04 | 0.06 | -0.11 | -0.24 |
| 1 | -0.01 | -0.15 | -0.27 | 0.09 | -0.38 | -0.08 | 0.06 | 0.03 | -0.2 |

In all, TWL estimates for one percent and 0.2 percent annual chance events had little to no variation, regardless of screening approach or surge/wave prioritization ratios. Typical differences between TWL-CSS and TWL-FSS were found to vary roughly between 0 ft and 0.20 ft. The main reason for these similar results, at least in the case of Lake Michigan, is the considerable overlap between surge and wave events. In general, approximately 25 to 35 percent of the events are both high-ranked surge events and high-ranked wave runup events. Many of these dual threat events form part of the extreme tail of the distribution and serve as anchors, thus minimizing the effects of different event prioritization ratios. Thus, the impact of selecting a 50 percent surge/50 percent waves POT approach versus a surge-only POT approach is minimized.

# 14 Final Storm Selection for Flood Hazard Mapping

Storm surge estimated from measured water levels, along with waves from wind/wave-surrogate analysis and hindcast waves are used to identify a large number of events around the lake using the peaks-over-threshold (POT) method. This method recognizes that total water level along the coast is generally a function of lake level, storm surge and wave runup components. Wave runup can dominate for steep shorelines or structures. If significant shore-fast ice is present, then wave runup will not occur. Thus selecting storms is not as simple as selecting the highest storm surge POT values. The analysis must properly weight the influence of surge, waves, and ice to rank and select the storms.

The CSS that will be used for the flood hazard mapping production consists of the top 150 storm events with, approximately, a 50 percent surge/50 percent wave prioritization ratio. The final storm selection process that was adopted for Lake Michigan is summarized below.

## 14.1 Initial identification of composite storm set

1. Identify storms having the highest peak storm surge at each of the long-term water level measurement sites using monthly maxima (highest hourly each month) and monthly average, and/or hourly water level data. Rank storms based on magnitude of peak surge.
2. Identify storms based on surrogate wave calculations made for each long-term water level measurement location. Used as inputs to the wave height calculation are: wind speed, wind direction, and fetch for that direction. These wind parameters are spatially interpolated at each of the long-term water level measurement sites using data available from nearby long-term meteorological stations; no ice cover is assumed. Rank storms based on magnitude of calculated peak significant wave height.
3. Starting with highest 20 ranked surge events at each location, define highest ranked surge events for the lake. If storms are *duplicates* (i.e. a storm is a large event at more than one site), reject duplicate event (or events) with the lowest site-specific rank and include the next largest storm surge event at that site. In general, the replacement procedure might involve going deeper in the rankings and sampling storms below the top 20.

4. Starting with highest 20 ranked wave events at each location, define highest ranked wave events for the lake; handle duplicates in the same way as for storm surge.
5. Balance the number of storms selected for each site to maximize consistency in geographical coverage of selected storms; define an initial set of 150 storms (roughly 50 percent based on maximum wave height and 50 percent based on peak storm surge).
6. Check for duplicate storms again by comparing surge events versus wave events. If duplicate storms are found among both sets, drop the storm with the lowest ranking in its corresponding subset (surge or waves) and follow the same replacement procedure discussed above, until the set of 150 unique storms is achieved.
7. Compute number of events that are wave events, surge events, and both wave/surge events. At this point the initial 150-storm set should consist roughly of 75 unique surge events and 75 unique wave events.

## 14.2 Verification, ice-screening, and finalization of the storm set

1. Break the lake shoreline into regions or segments (e.g., for Green Bay: north, south, east, and west coastline segments; and for main Lake Michigan: north, northeast, southeast, south, southwest, and northwest segments).
2. Using the ice maps nearest the time of each of the 150 storms, or the ice maps that are bracketing the storm, record if shore-fast ice is present in each of the various coastline segments of the lake , and either partially or fully blocking waves in that segment (greater than 70 percent concentration assumed to be blocking waves).
3. For each of the 150 storms, decide if any wave events defined using the surrogate calculation process (which did not consider ice cover) should be removed because of prevalent shore-fast ice conditions during that storm. Identify candidate storms to remove from the set and note which segment(s) are affected by the removal.
4. If a storm in the initial set is a low-ranked surge event and does not appear to be a large wave event, consider it as a candidate to remove.
5. Examine WIS wave data nearest each site, if available, or any other storm data sources (such as extreme winds favorable to surge/wave generation); rank the events, check to see if top 15 events in WIS wave data set are included in the initial storm set defined using the surrogate wave calculation approach.

6. Identify events, particularly large wave events, not included in the initial storm set as candidate storms to add. Examine the peak surge and wind conditions during these events.
7. Examine low surge events or surrogate wave events that had prevalent shore fast ice coverage and replace them with wave events defined from the analysis of WIS wave data.

For the examination of events for Lake Michigan, these checks resulted in replacement of 20 storms that were contained in the original 150-storm set. The final list of 150 storms selected as the CSS for Lake Michigan Flood Hazard Mapping production is presented in Appendix C.

## 14.3 Additional statistical-analysis topics

The following is a summary of additional issues that arose during discussions of the applicability of the storm selection and statistical analysis approach.

### 14.3.1 Sampling on varying lake levels

Flooding events primarily occur on high lake levels. The sampling technique employed as part of the statistical methodology picks storm events across all lake levels. As shown in Chapter 8 (*Sampling Across High and Low Lake Levels*), all high lake levels periods are adequately represented in the sampling.

### 14.3.2 Screening/sampling fraction

Analysis showed that a 50 percent/50 percent POT sampling of surge/waves was reasonable. As shown in Chapter 12, *Surge/wave event prioritization* Section, using 50 percent surge/50 percent waves or 100 percent surge/0 percent waves fractional sampling, did not impact the TWL distributions in a statistically significant way. The final storm list may vary based on the screening method, but the impact on flood mapping will not be significant, at least for Lake Michigan. Typical differences in extreme one percent and 0.2 percent exceedance TWL values for these two methods were in the range of 0. 10 to 0.20 ft.

### 14.3.3 Summer convective events

Presently available wind field data and wind estimation techniques do not allow accurate modeling of convective events. As discussed in Chapter 12,

the effect of including or not including summer events (convective storms) has no significant impact on the TWL distribution. The recommended approach consists of modeling events that occur during the May-September period only if they are well organized low pressure systems, and to ignore all fast moving disorganized convective systems (thunderstorm/squall line). The recommended approach is to list all of the top events that occurred during the summer period, and then analyze storm surge time series, meteorological data, and weather journals to screen out the convective systems.

### 14.3.4 Ill-formed storm surge distributions

Storm surge elevation return period distributions can be ill-formed (concave up curves) primarily because the most extreme probabilities of surge events can be difficult to determine because they are a function of data record length. However, TWL distributions, which by definition include runup, appear to be well formed and concave down. The FEMA flood hazard mapping effort focuses on TWL, rather than surge-only elevations, to establish BFEs.

### 14.3.5 Poorly fitting TWL distributions and anti-storms

Composite POT storm sampling emphasizes low frequency events (extreme values), thus resulting in some poor fits in the high-frequency portion of the distribution and potential underestimation of the bulk of the distributions. Also, storms selected as high surge at one end of a long lake often are observed as lower lake levels at the other end of the lake. Consequently, the high-frequency end of the TWL distributions could potentially be loaded with artificially low surge values (i.e., anti-storms).

To solve these problems, an optimization technique for fitting the distribution was developed based on Q-Q plots, as discussed in Chapter 10 (SWL – FSS). This method optimizes GPD fits by focusing on the low-frequency end of the distributions. Storms not part of the extreme events set (e.g., low surge/water level conditions, or anti-storm) are dropped from the distributions using this process.

### 14.3.6 Sampling window size

In the case of Lake Michigan, clearly i.i.d. storms occur at least 96 hours apart. On the other hand, a single storm event could take up to 12 hours to

translate from one end of an elongated lake to the other. Therefore, a time "buffer" of 24-48 total hours (equivalent to 12-24 hours on either side of peak) was found to be large enough to adequately identify duplicate storms among different gages, but small enough to clearly identify independent events. The approach recommended here uses a 24-hour sampling window on either side of the storm peak as part of the storm sampling technique.

### 14.3.7 Record length and data availability

Inadequate record length is a primary contributor to error in estimating BFEs and every effort should be made to extend the record length to include the entire 1960-2010 period because it has been shown that a 50-year record length is the minimum acceptable length. Storm screening for decades for which data, such as hourly water levels, are scarce (e.g., 1960s decade) can benefit from using wind data and WIS, or other hindcast results as guides.

# 15 Conclusions and Recommendations

This study focused on evaluating the validity of the statistical analysis and storm sampling approach recommended by Melby et al. (2012), and determining the adequate storm sample size. For this purpose, several tasks were performed, including: computation of storm-surrogate waves and water levels; definition of full sample and composite storm sets; and evaluation of the statistical analysis approach for a record length of 50 years.

Storm-surrogate analysis was used to estimate surge and wave heights, based on wind data, for the entire 1960-2010 period. A steady-state surge model did not have sufficient skill. Storm surge magnitudes were in the correct order of magnitude, but their accuracy was not adequate for use as part of storm screening and sampling process. Instead, measured water level data were used. Wave heights, however, were estimated using surrogate calculation methods with much better results. Wave estimates were compared and validated against WIS hindcast data for the 1970-1997 period. The results showed that, for Lake Michigan, RMS deviations of surrogate waves vs. WIS hindcasts are typically less than 20 percent. It is concluded that the use of the ACES/CEM wind-wave generation methodology is an adequate alternative, in absence of measured or hindcast data, to estimate waves as part of the storm selection process.

The prioritization of waves and water levels in storm sampling was extensively evaluated. When developing the full storm sets (FSS) and the composite storm set (CSS) it was found that approximately 25 percent to 30 percent of all selected storms were both high ranked surge events and high ranked wave events. This result is not surprising since both waves and surge are forced by wind. In other words, the types of storm events are not mutually exclusive; there is considerable overlap between high storm surge elevations and high-energy wave events. Many of these *dual threat* events form part of the extreme tail of the TWL distributions and serve as anchors, thus minimizing the effects of different event prioritization ratios. Extreme TWL and SWL values, corresponding to one percent and 0.2 percent annual chance of exceedance, had negligible variation regardless of the event prioritization ratio. Similarly it was observed that the elimination of

summer convective events did not produce any adverse effects or significant differences in the TWL distributions.

The statistical analysis approach originally recommended by Melby et al. (2012), which was used to assess a 27-year record length as part of the Lake Michigan Pilot Study was successfully applied, relatively unchanged, to the 50-year period from 1960 to 2010. Still, several conclusions were reached and improvements were made as consequences of this focus study and discussions among technical team members during its execution. It was established that the ideal number of events that should be sampled, to get accurate estimates of the extreme events and the correct shapes of the water level distributions, is approximately 150 storms. A smaller number of events (e.g., 100 storms) produced adequate water level estimates, but the shapes of the distributions were incorrect, particularly in the high-frequency portion of the distributions. Sampling more than 150 storms did not result in any appreciable accuracy improvements.

The Q-Q plot optimization technique introduced here proved to be a very useful tool to obtain the best possible GPD fits for the water level data. The fit optimization eliminates non-extreme events, or even anti-storms, that can potentially be introduced by the CSS approach; only events that are confirmed to be part of the same extreme event population are retained. This technique improves the overall distribution shape and the accuracy of the extreme water level estimates.

Also, the issue of sampling across all lake levels was further assessed through resampling analysis. It was shown that the 150-storm CSS not only captures all decadal variation in water levels but also much of the higher frequency variation in lake-wide water levels. Therefore, all high lake levels periods are adequately represented.

# References

Cardone, V.J. 1969. Specification of the Wind Distribution in the Marine Boundary Layer for Wave Forecasting. Tech. Rept. 69-1, Geophys. Sci. Lab., New York Univ., New York, N.Y.

Dean, R.G., and R.A. Dalrymple. 2001. Coastal Processes with Engineering Applications. Cambridge University Press.

Dean, R.G., and R.A. Dalrymple. 1990. Water Wave Mechanics for Engineers and Scientists. World Scientific Publishing Company.

Donelan, M. A. 1980. Similarity Theory Applied to the Forecasting of Wave Heights, Periods and Directions. Proc. Canadian Coastal Conf., Canada, National Research Council, pp 47–61.

Leenkenecht, D.A., A. Szuwalski, and A.R. Sherlock. 1992. Automated Coastal Engineering System (ACES): Technical Reference. US Army Engineer Research and Development Center, Vicksburg, MS.

Mase, H. 1989. Random Wave Runup Height on Gentle Slope. Journal of Waterway, Port, Coastal, and Ocean Engineering, Vol. 115, No. 5, pp 649-661.

Mathworks 2010. Surface Fitting Using *gridfit.m*. http://www.mathworks.com/matlabcentral/fileexchange/8998"; downloaded on July 28, 2010.

Melby, J.A., N.C. Nadal-Caraballo, and B.A. Ebersole. 2012. Wave Height and Water Level Variability on Lake Michigan. Draft Technical Report; US Army Engineer Research and Development Center, Vicksburg, MS.

Resio, D.T., and C.L. Vincent. 1977. Estimation of Winds over the Great Lakes. Amer. Soc. Civil Eng.Waterway, Port, and Coast. Ocean. Div. J., No. 102, pp 265-283.

Schwab, D.J. 1978. Simulation and Forecasting of Lake Erie Storm Surges. Mon. Wea. Rev., No. 106, pp. 1476-1487.

Schwab, D.J., and J.A. Morton, J.A. 1984. Estimation of Overlake Wind Speed from Overland Wind Speed: A Comparison of Three Methods. J. Great Lakes Res. 10(1), pp 68-72.

Smith, J. M. 1991. Wind-Wave Generation on Restricted Fetches. Misc. Paper CERC-91-2, Coast Engineering Research Center, U.S. Army Eng. Wtrwy Experiment Station Misc. Paper CERC-91-2, 43 pp.

Stockdon, H.F., R.A. Holman, P.A. Howd, and A.H. Sallenger. 2006. Empirical Parameterization of Setup, Swash, and Runup. Coastal Engineering 53, Elsevier, pp 573-588.

Van Dorn, W.C. 1953. Wind Stress on an Artificial Pond. J. Mar. Res., Vol. 12.

USACE. 2002. Coastal Engineering Manual – Chapter 2: Meteorology and Wave Climate. Engineering Manual 1110-2-1100, U.S. Army Corps of Engineers, Washington, DC.

USACE. 1984. Shore Protection Manual. 4th Ed., Waterways Experiment Station, U.S. Government Printing Office, Washington, DC.

# Appendix A: Q-Q plots and CDF for TWL-FSS

This appendix includes the Q-Q plots and CDF plots derived from TWL-FSS corresponding to the Ludington and Milwaukee gages, as discussed in Chapter 12 (TWL – FSS):

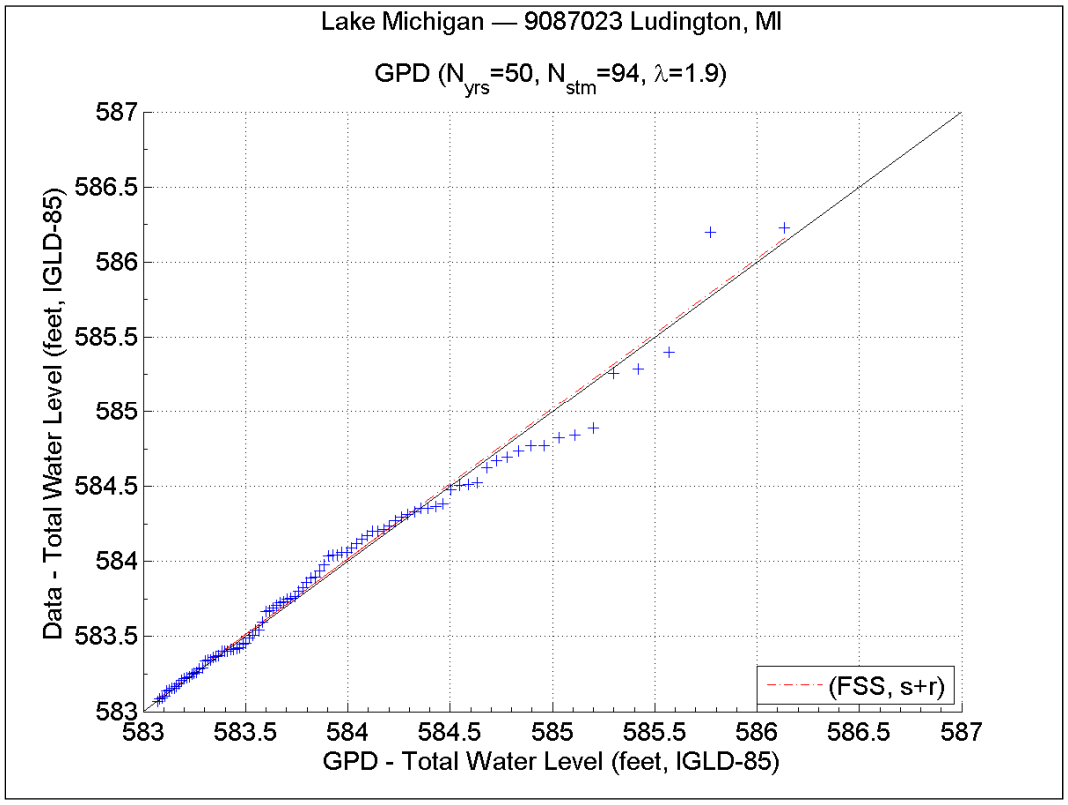



Figure A1. TWL Q-Q plot from FSS (s+r, with convective storms) for Ludington, MI.

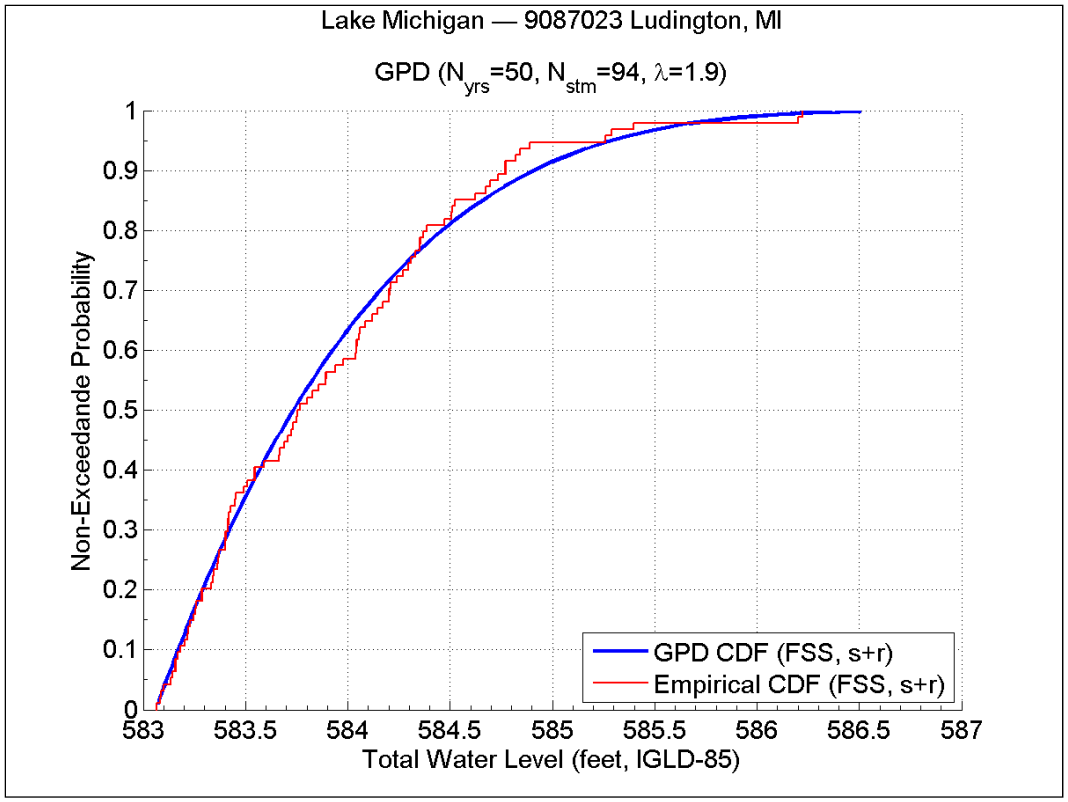



Figure A2. TWL CDF plot from FSS (s+r, with convective storms) for Ludington, MI.

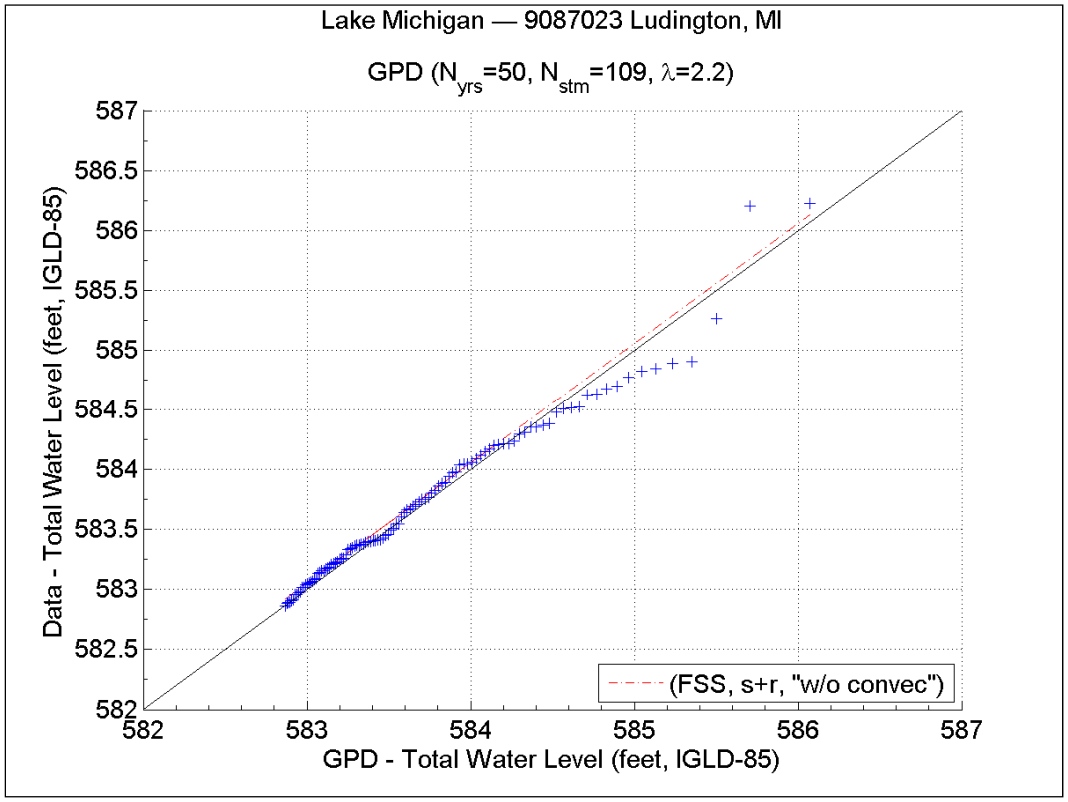



Figure A3. TWL Q-Q plot from FSS (s+r, no convective storms) for Ludington, MI.

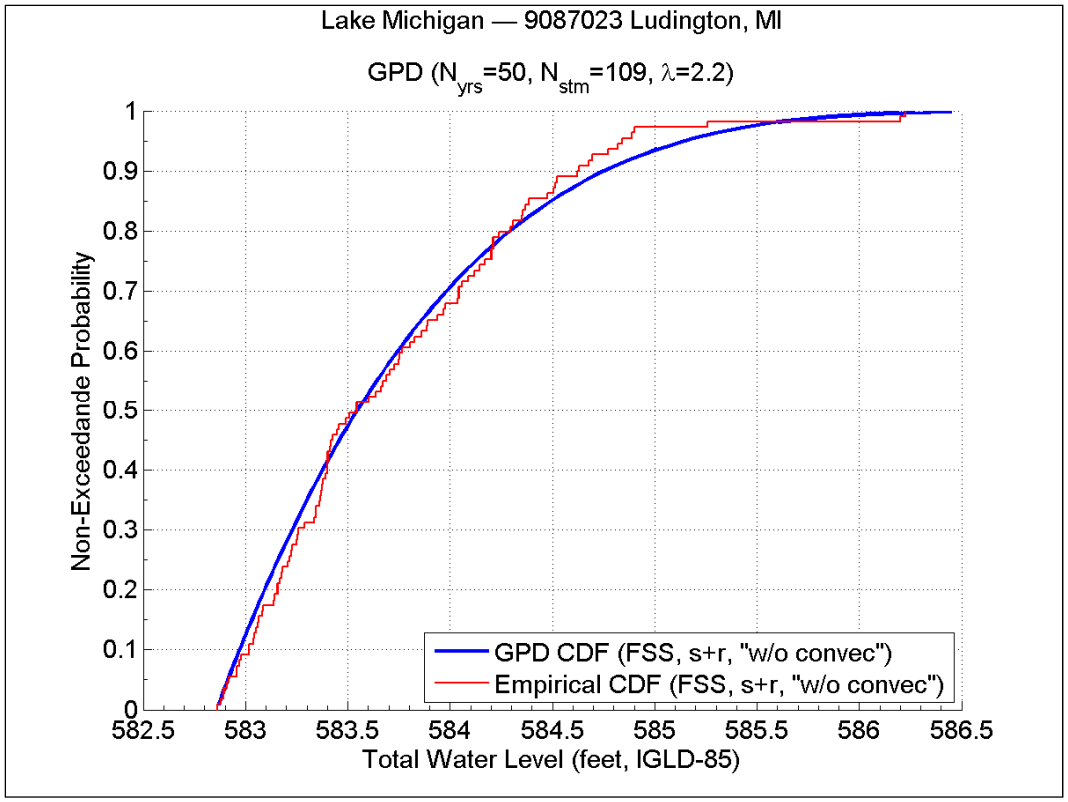



Figure A4. TWL CDF plot from FSS (s+r, no convective storms) for Ludington, MI.

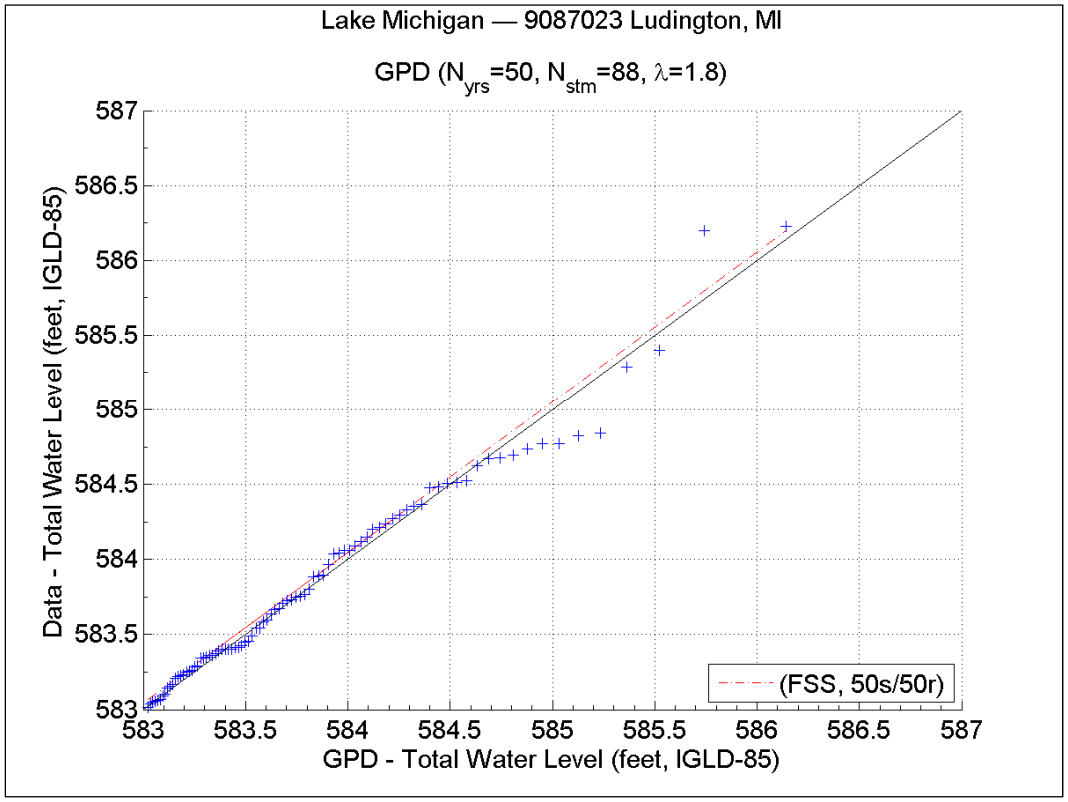



Figure A5. TWL Q-Q plot from FSS (50s/50r, with convective storms) for Ludington, MI.

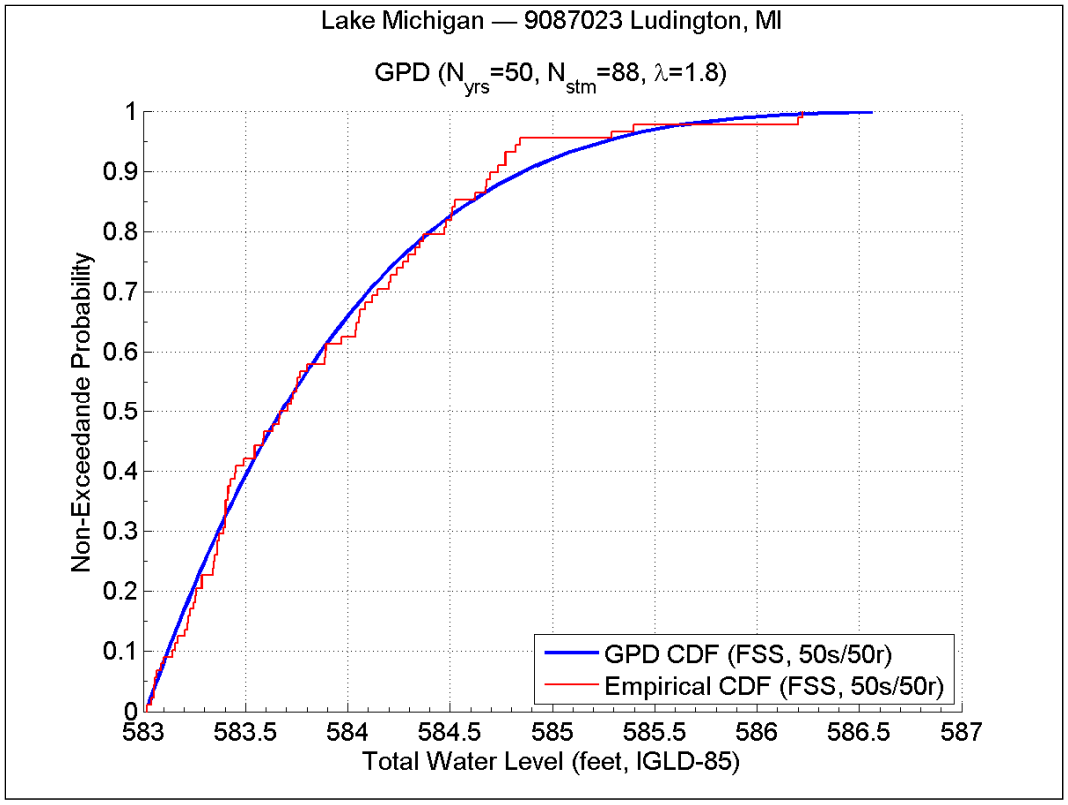



Figure A6. TWL CDF plot from FSS (50s/50r, with convective storms) for Ludington, MI.

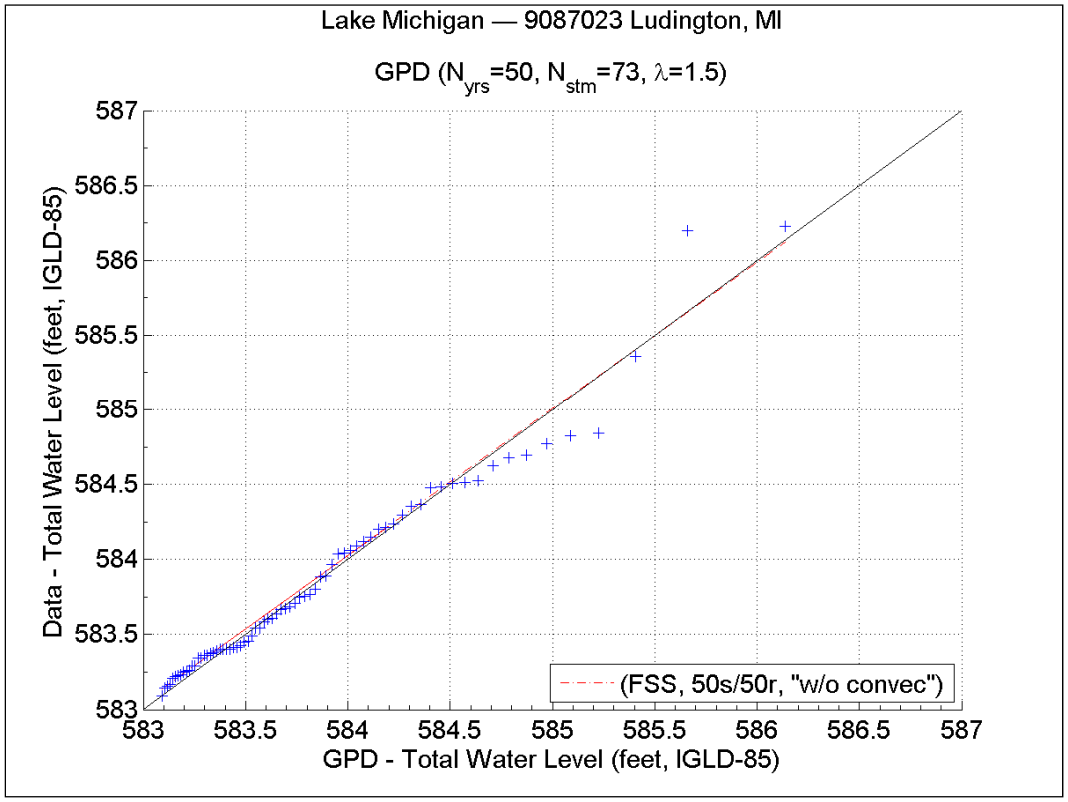



Figure A7. TWL Q-Q plot from FSS (50s/50r, no convective storms) for Ludington, MI.

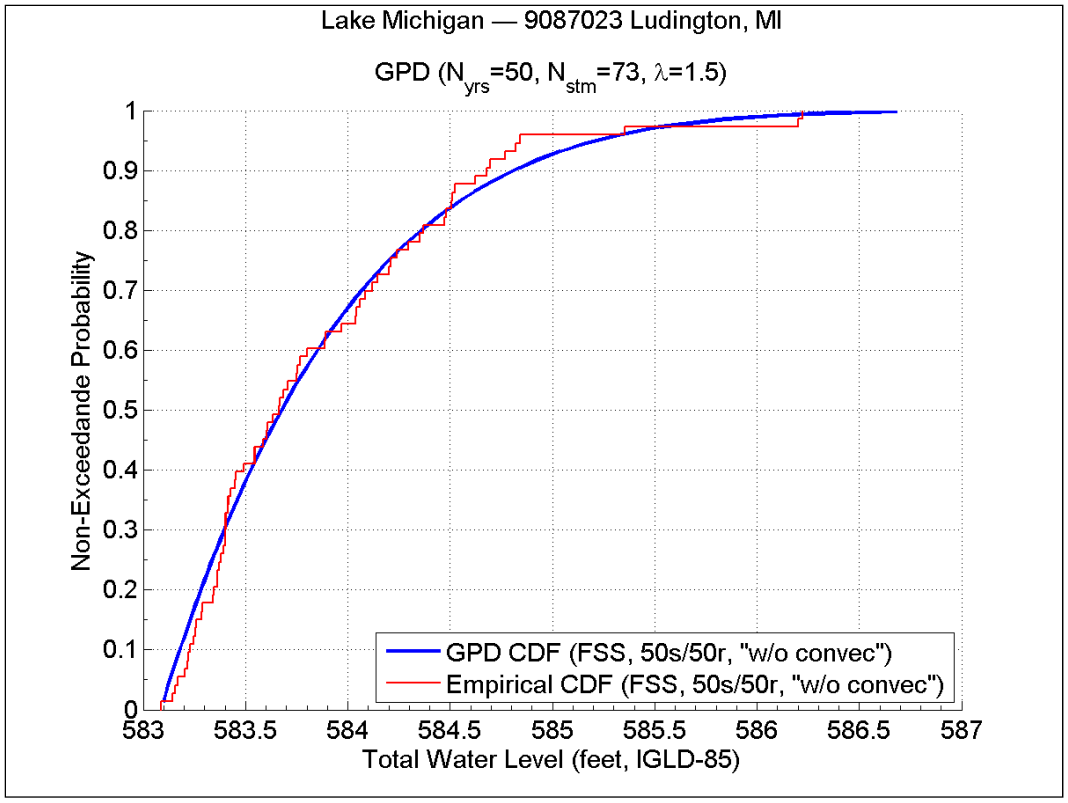



Figure A8. TWL CDF plot from FSS (50s/50r, no convective storms) for Ludington, MI.

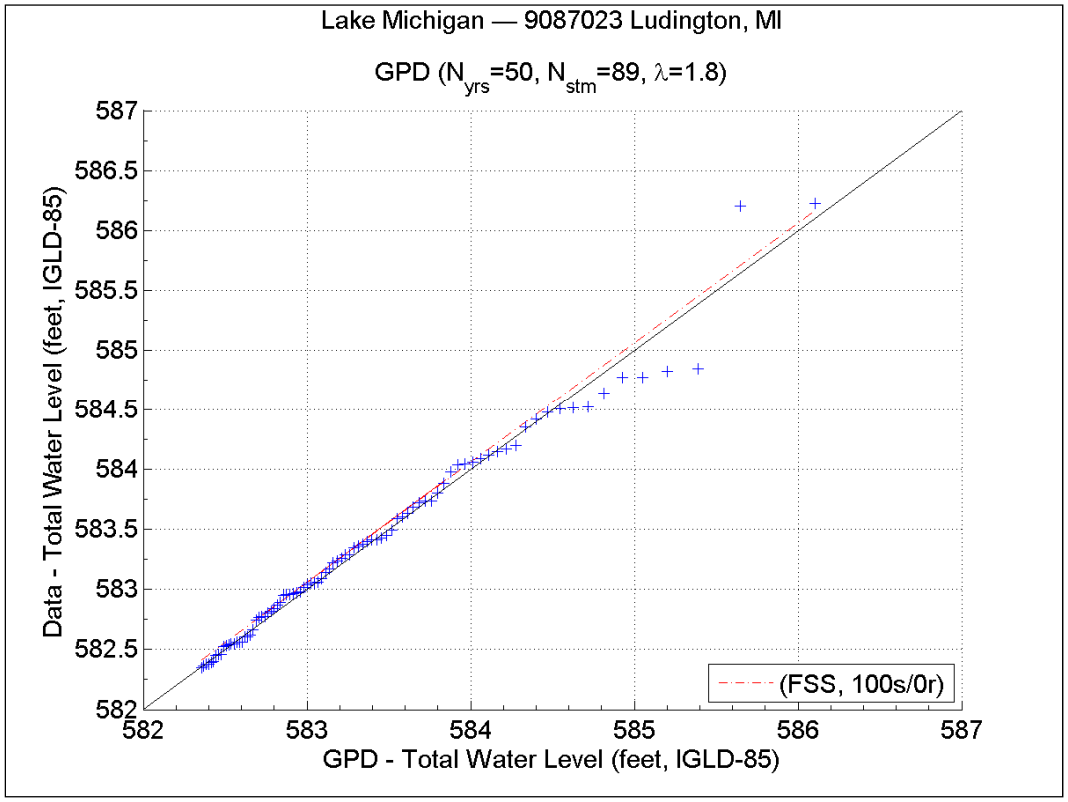



Figure A9. TWL Q-Q plot from FSS (100s/0r, with convective storms) for Ludington, MI.

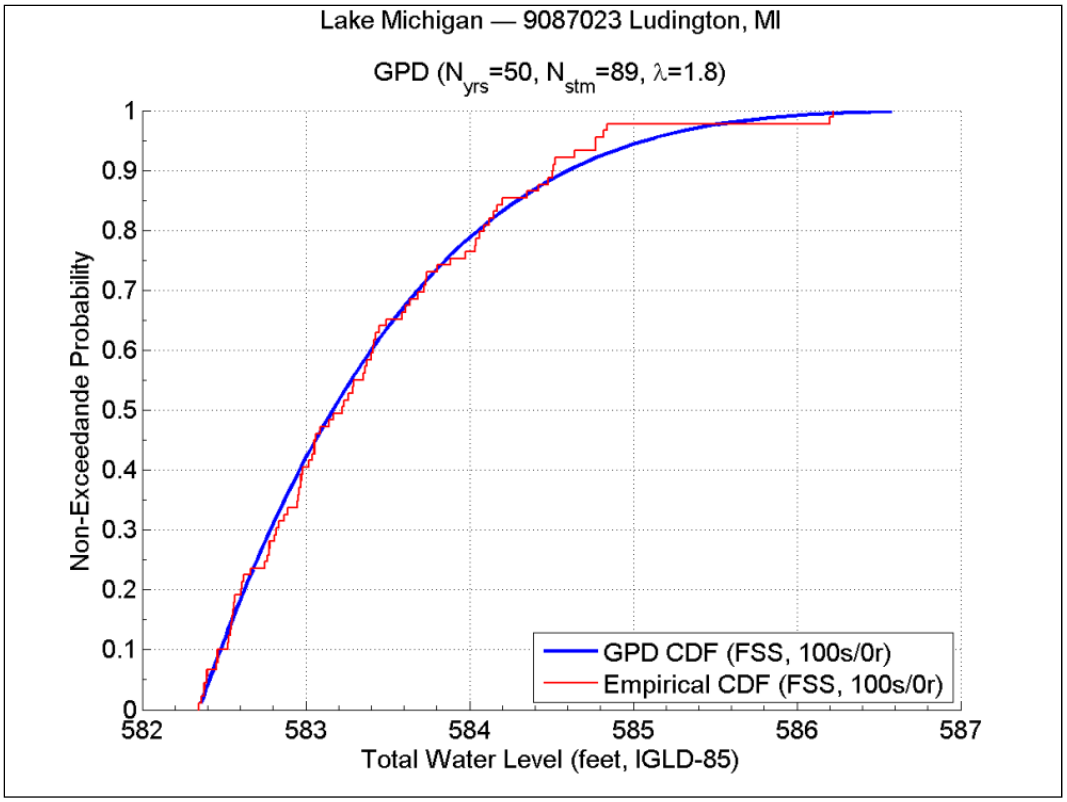



Figure A10. TWL CDF plot from FSS (100s/0r, with convective storms) for Ludington, MI.

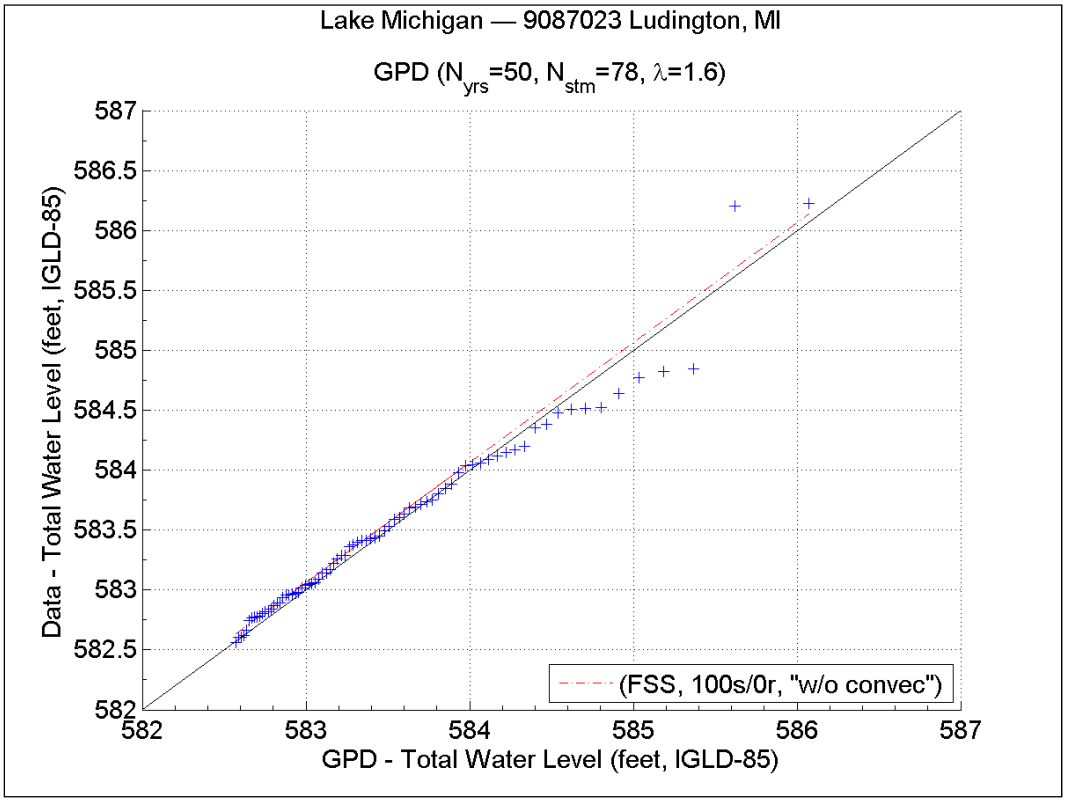



Figure A11. TWL Q-Q plot from FSS (100s/0r, no convective storms) for Ludington, MI.

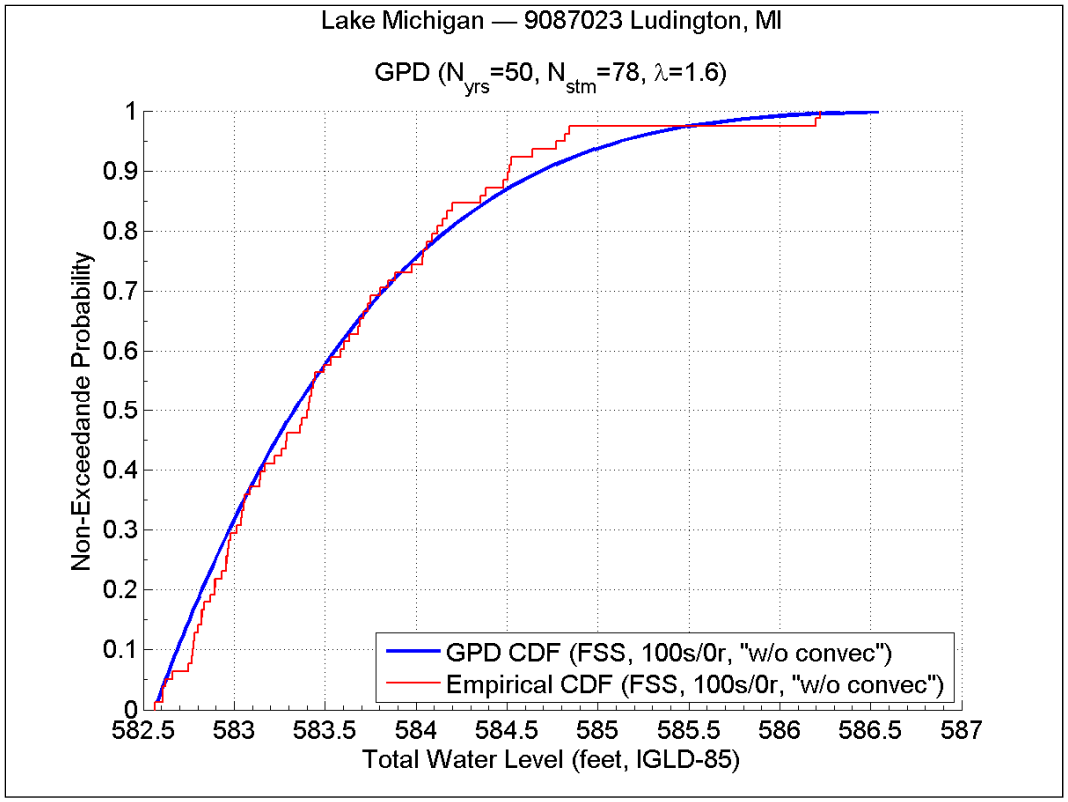



Figure A12. TWL CDF plot from FSS (100s/0r, no convective storms) for Ludington, MI.

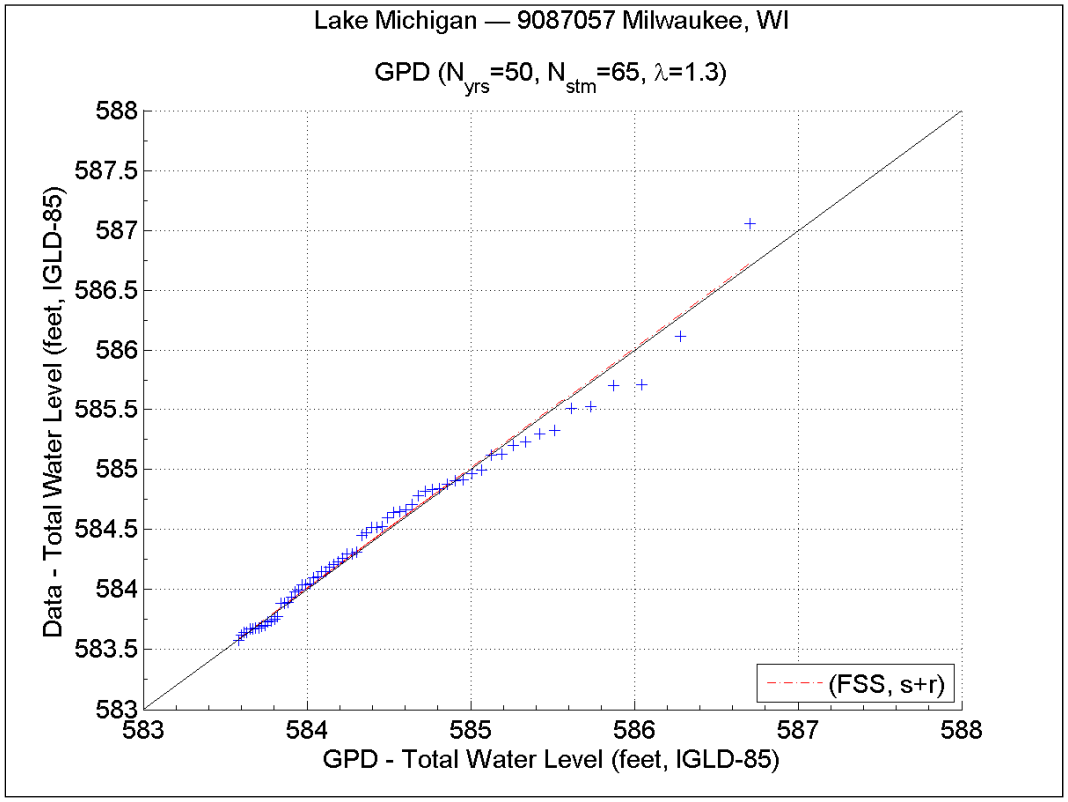



Figure A13. TWL Q-Q plot from FSS (s+r, with convective storms) for Milwaukee, WI.

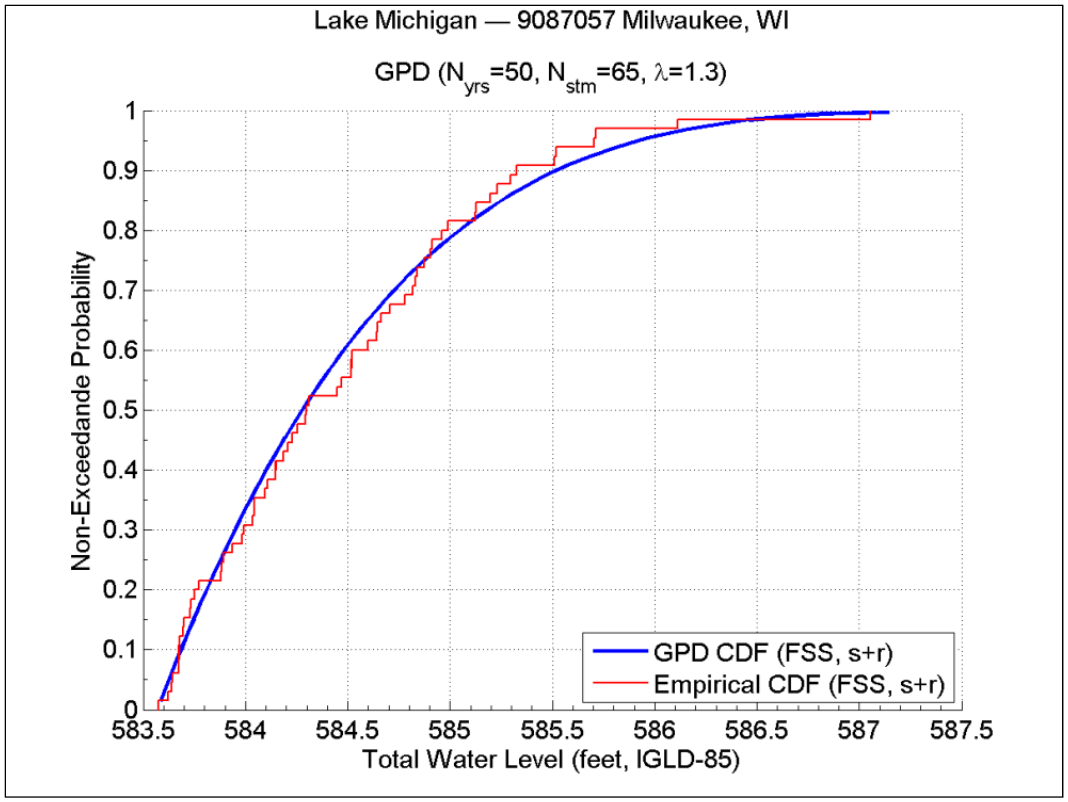



Figure A14. TWL CDF plot from FSS (s+r, with convective storms) for Milwaukee, WI.

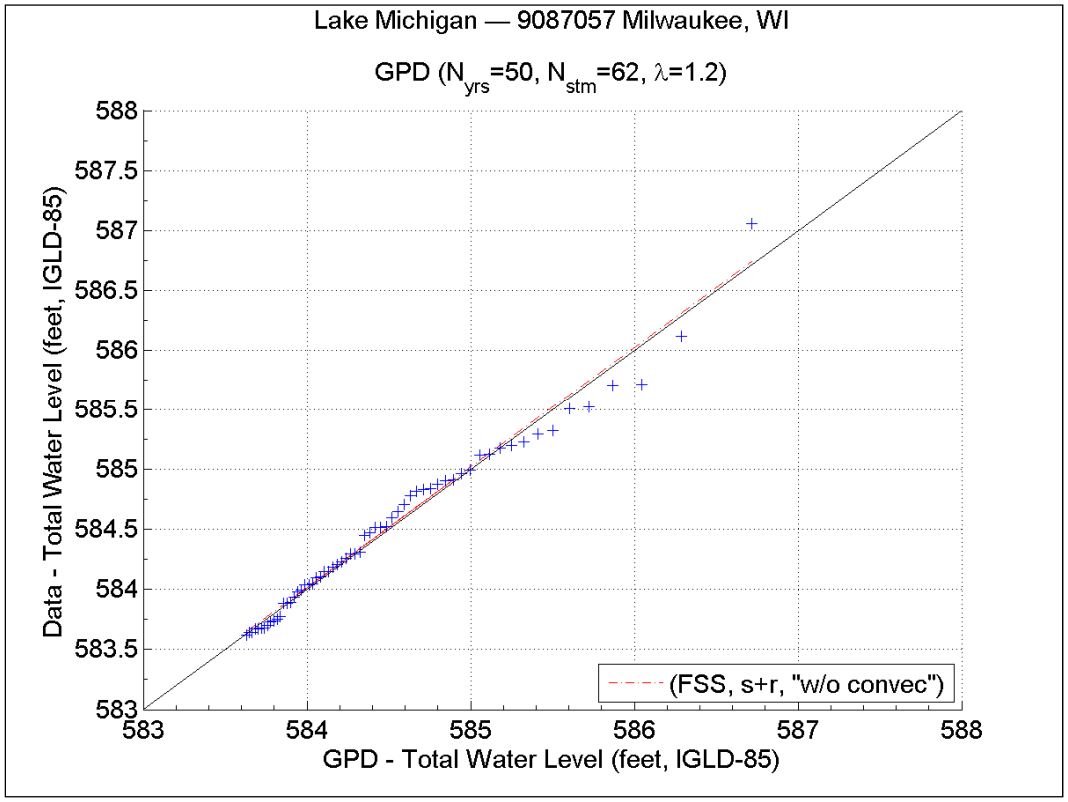



Figure A15. TWL Q-Q plot from FSS (s+r, no convective storms) for Milwaukee, WI.

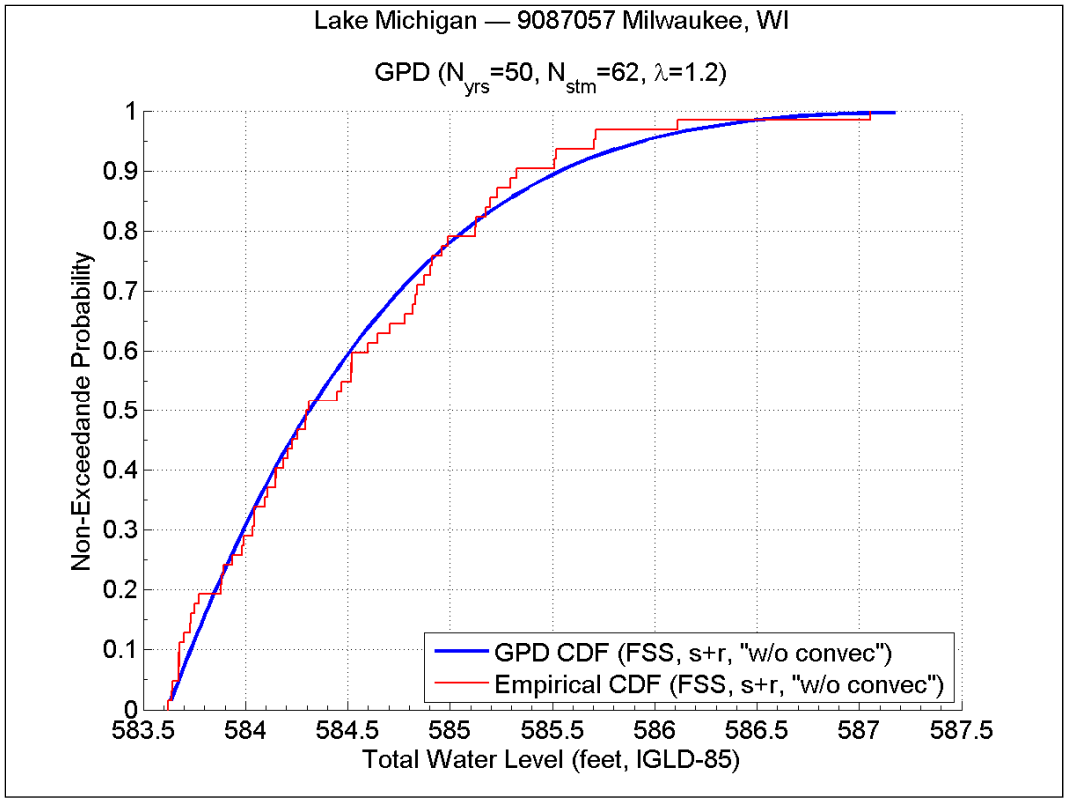



Figure A16. TWL CDF plot from FSS (s+r, no convective storms) for Milwaukee, WI.

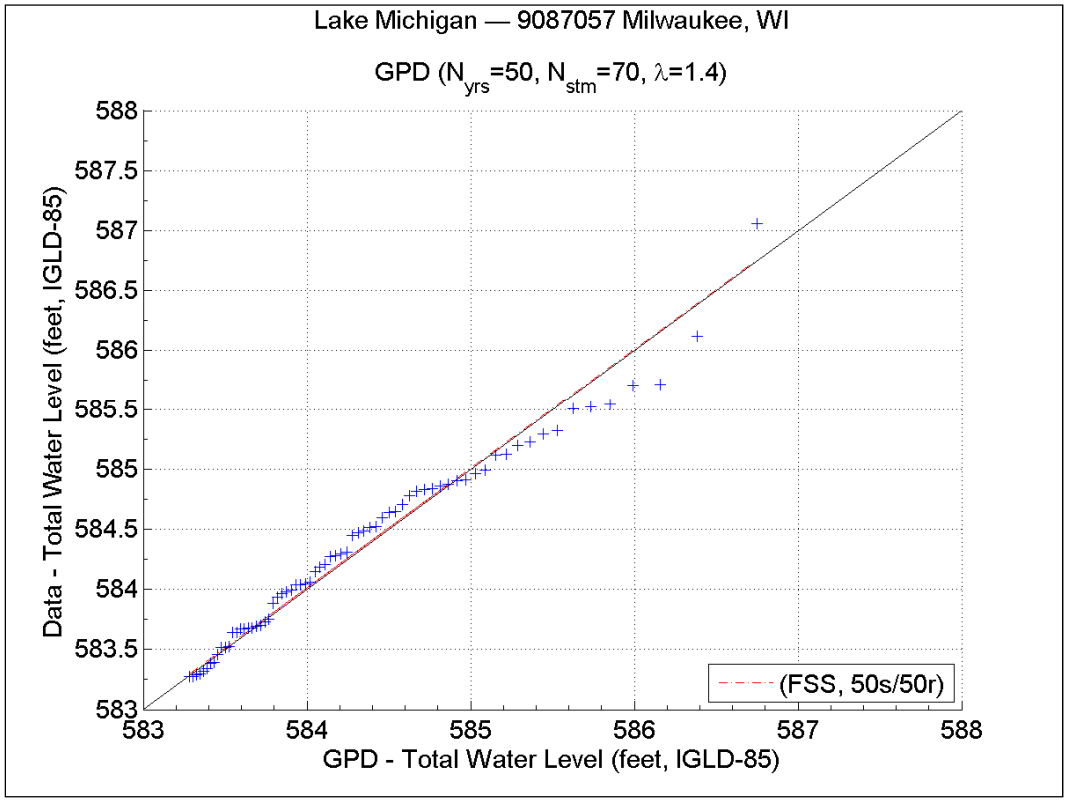



Figure A17. TWL Q-Q plot from FSS (50s/50r, with convective storms) for Milwaukee, WI.

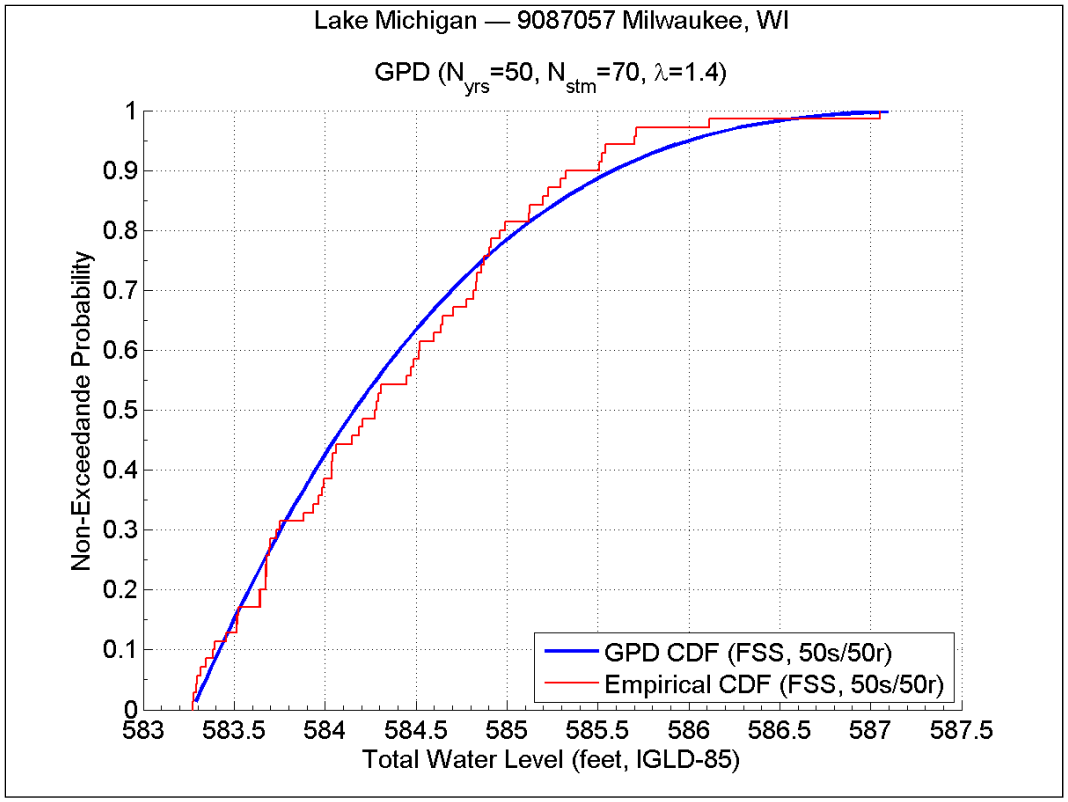



Figure A18. TWL CDF plot from FSS (50s/50r, with convective storms) for Milwaukee, WI.

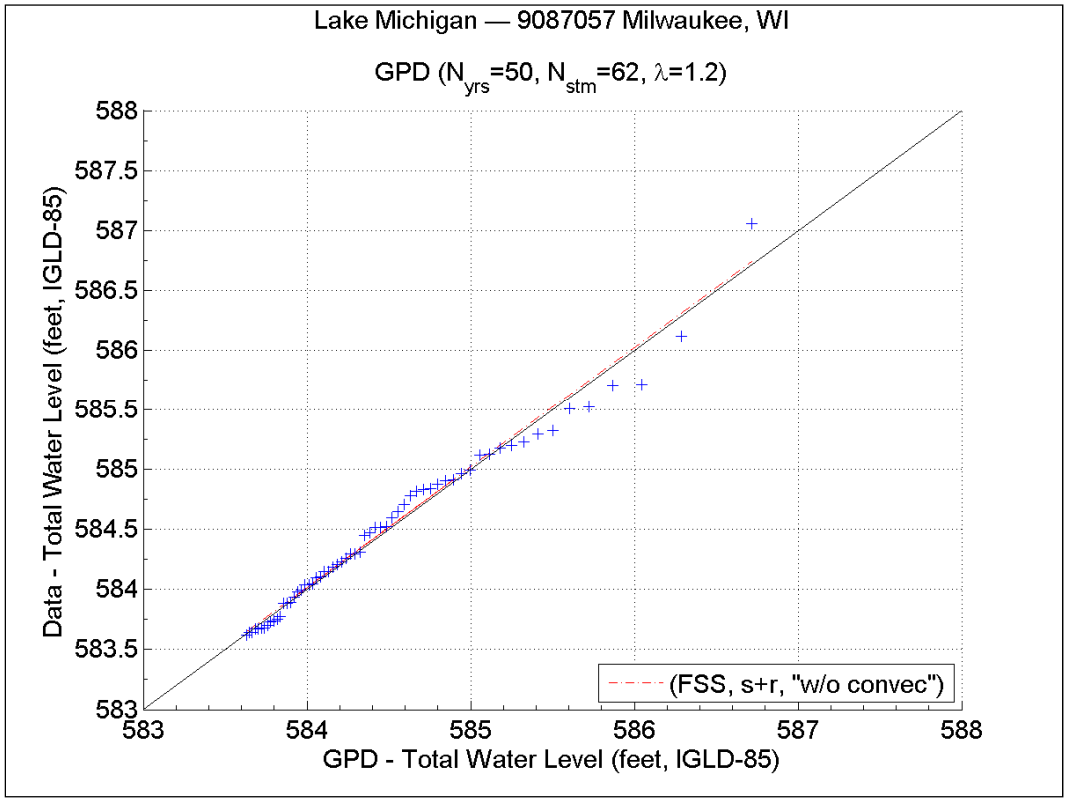



Figure A19. TWL Q-Q plot from FSS (50s/50r, no convective storms) for Milwaukee, WI.

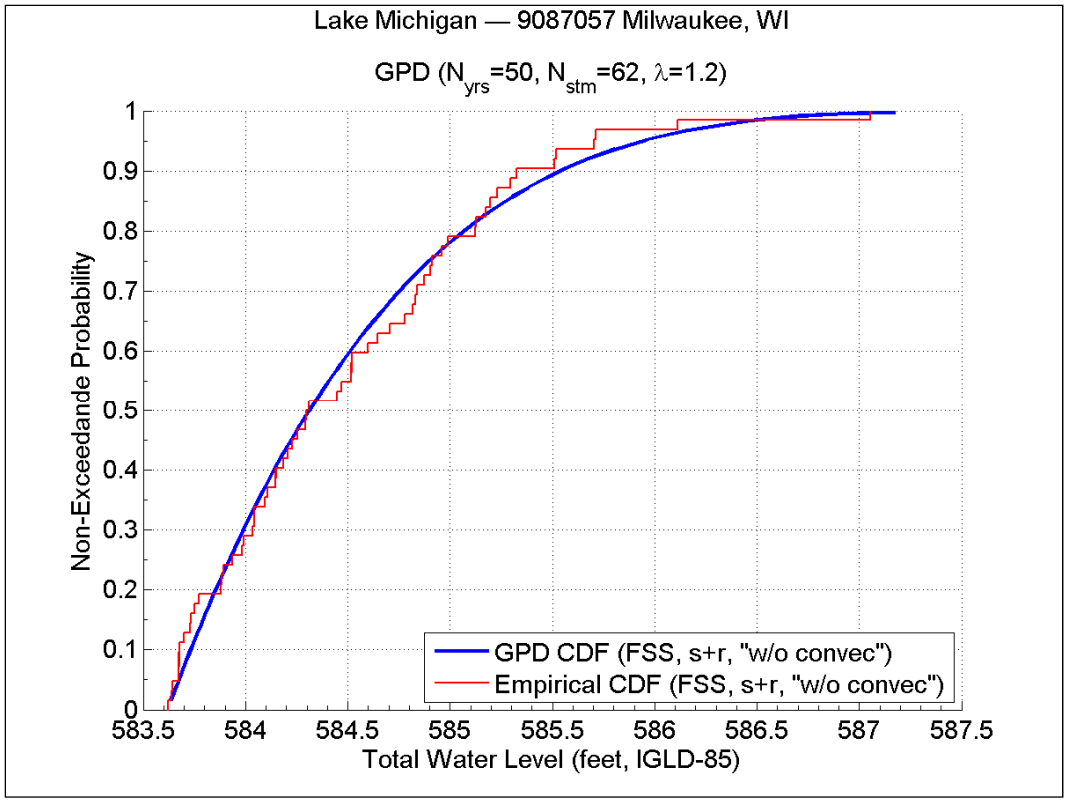



Figure A20. TWL CDF plot from FSS (50s/50r, no convective storms) for Milwaukee, WI.

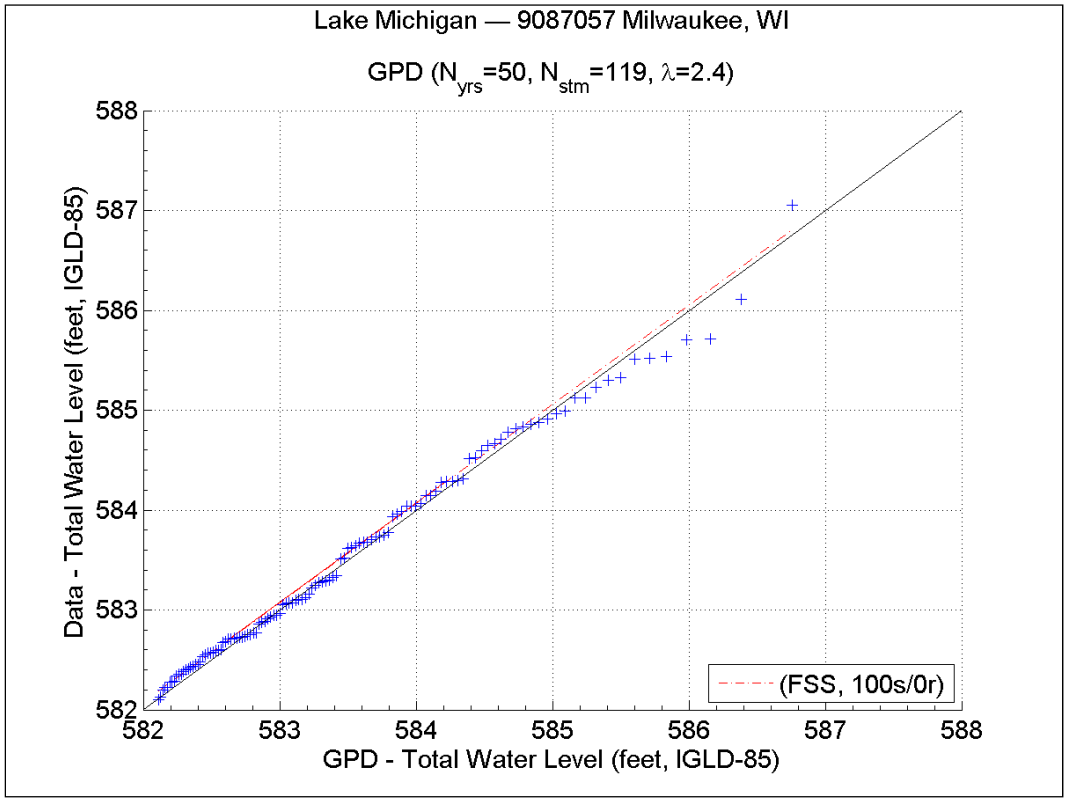



Figure A21. TWL Q-Q plot from FSS (100s/0r, with convective storms) for Milwaukee, WI.

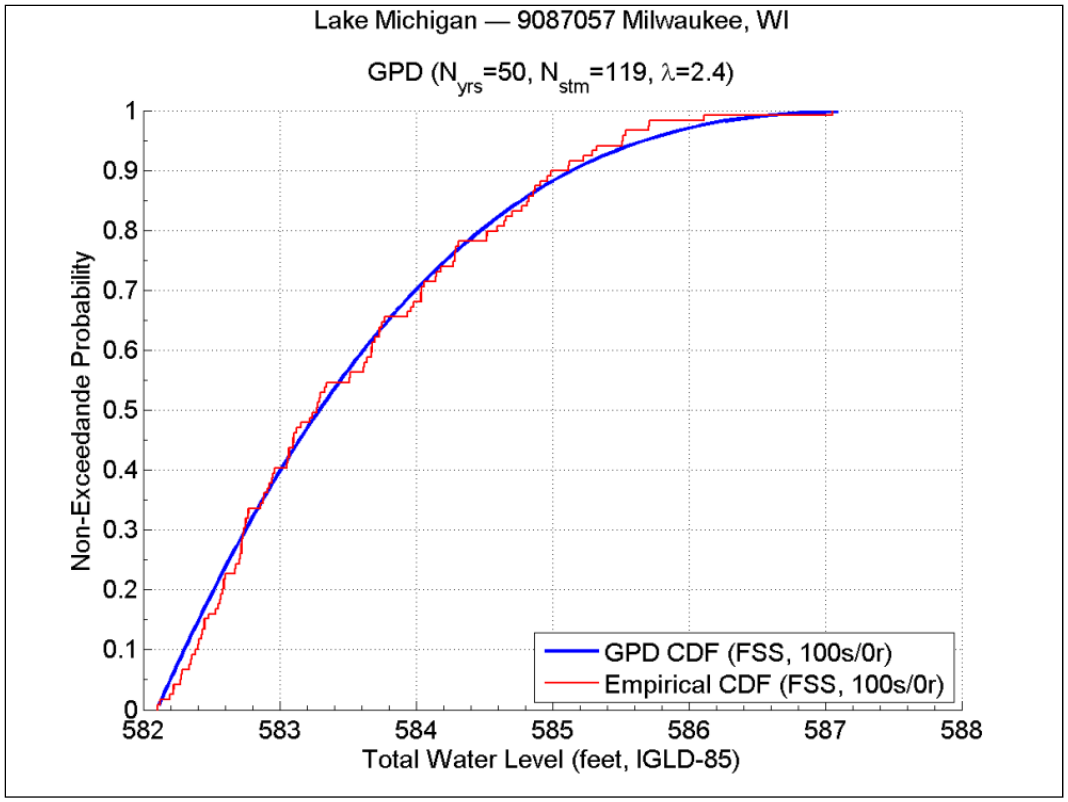



Figure A22. TWL CDF plot from FSS (100s/0r, with convective storms) for Milwaukee, WI.

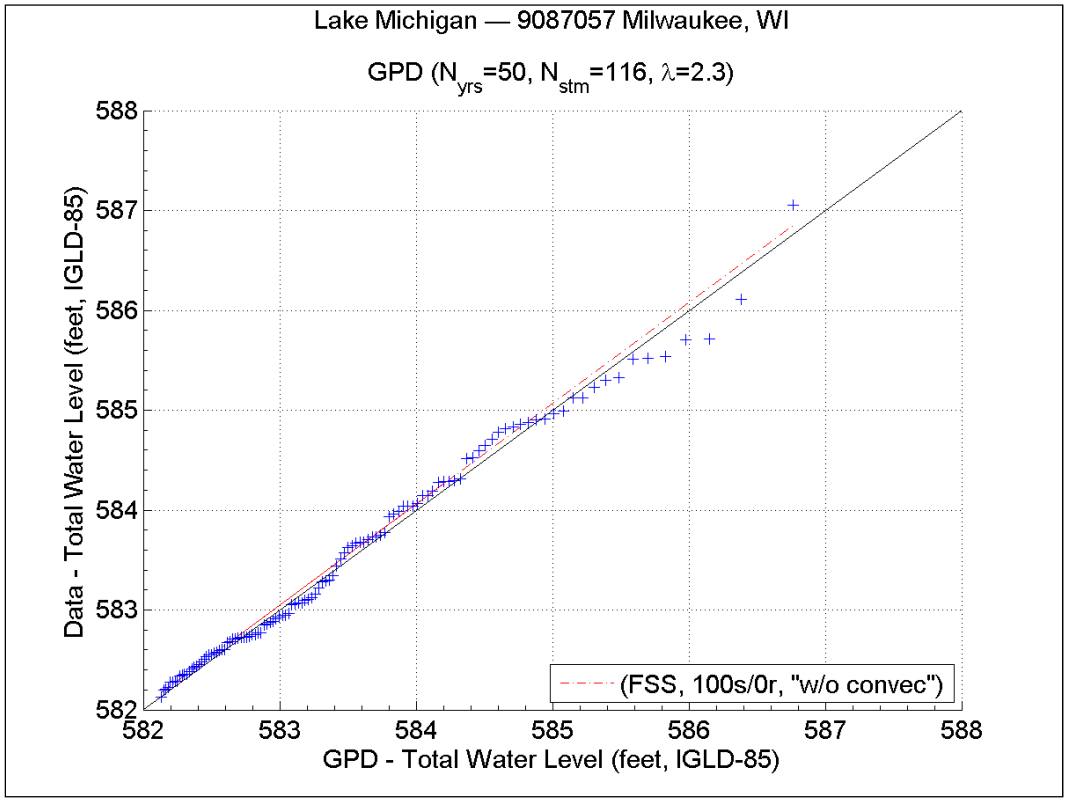



Figure A23. TWL Q-Q plot from FSS (100s/0r, no convective storms) for Milwaukee, WI.

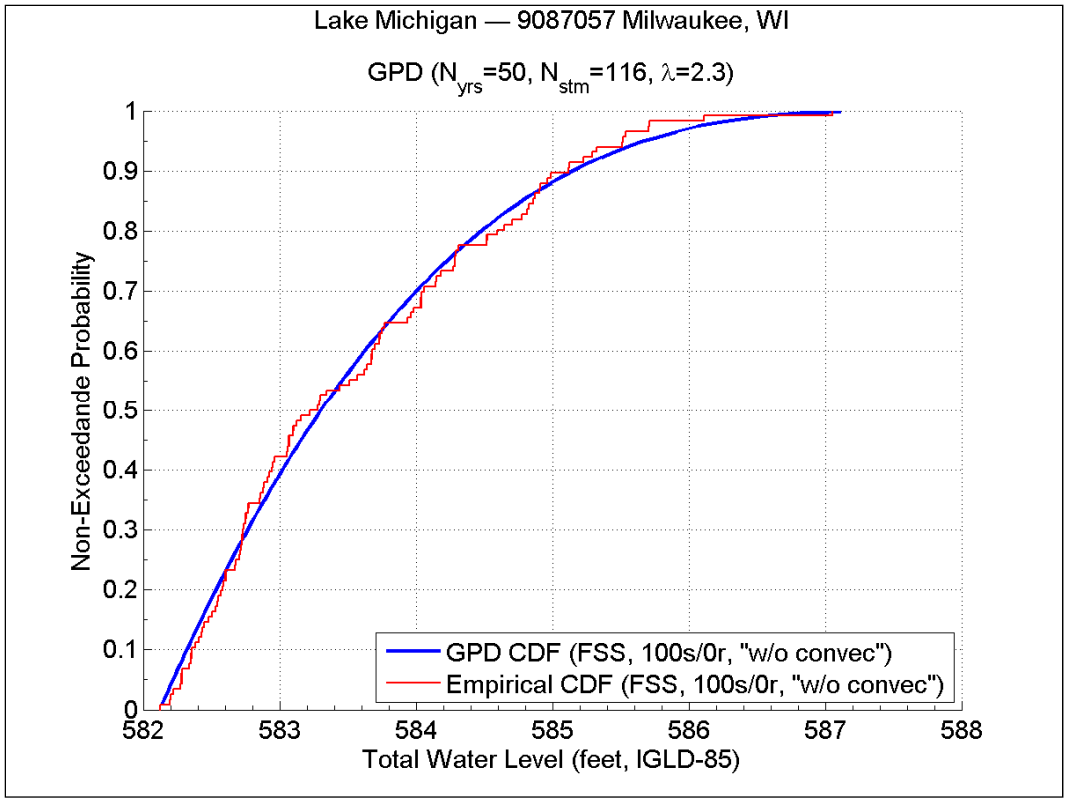



Figure A24. TWL CDF plot from FSS (100s/0r, no convective storms) for Milwaukee, WI.

# Appendix B: Q-Q plots and CDF for TWL-CSS

Following are the Q-Q plots and CDF plots from TWL-CSS corresponding to the Ludington and Milwaukee gages, as discussed in Chapter 13 (TWL – CSS):

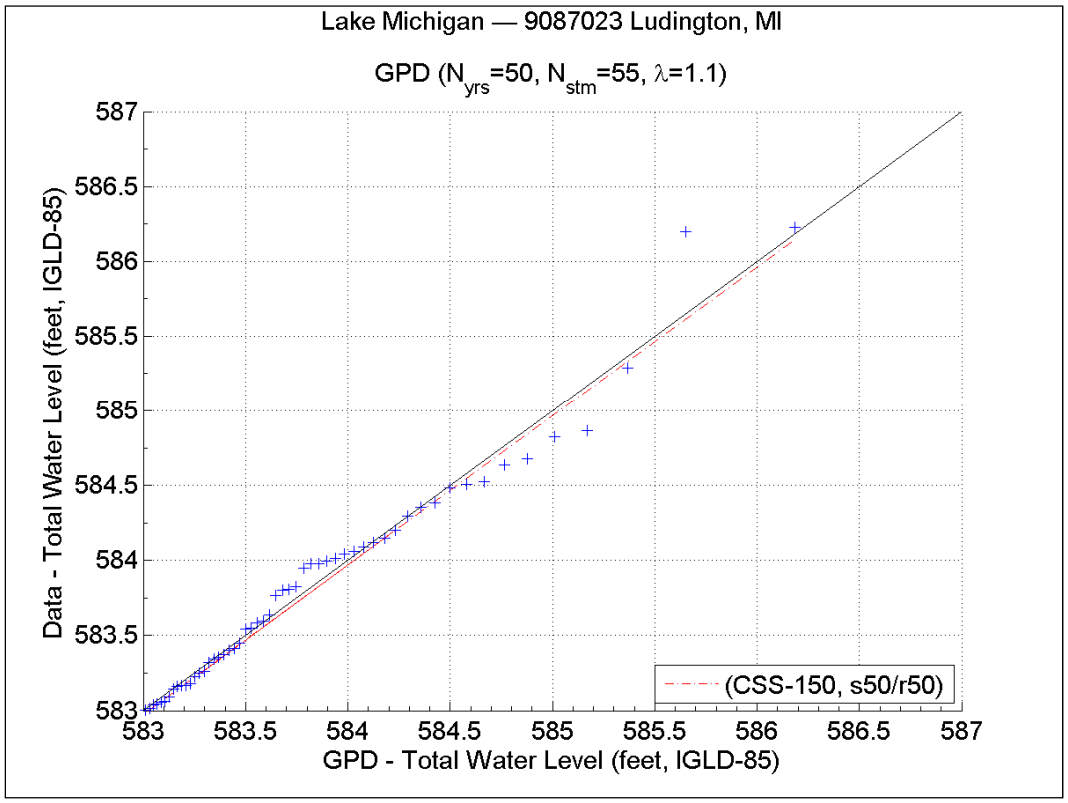



Figure B1. TWL Q-Q plot from CSS (50s/50r, with convective storms) for Ludington, MI.

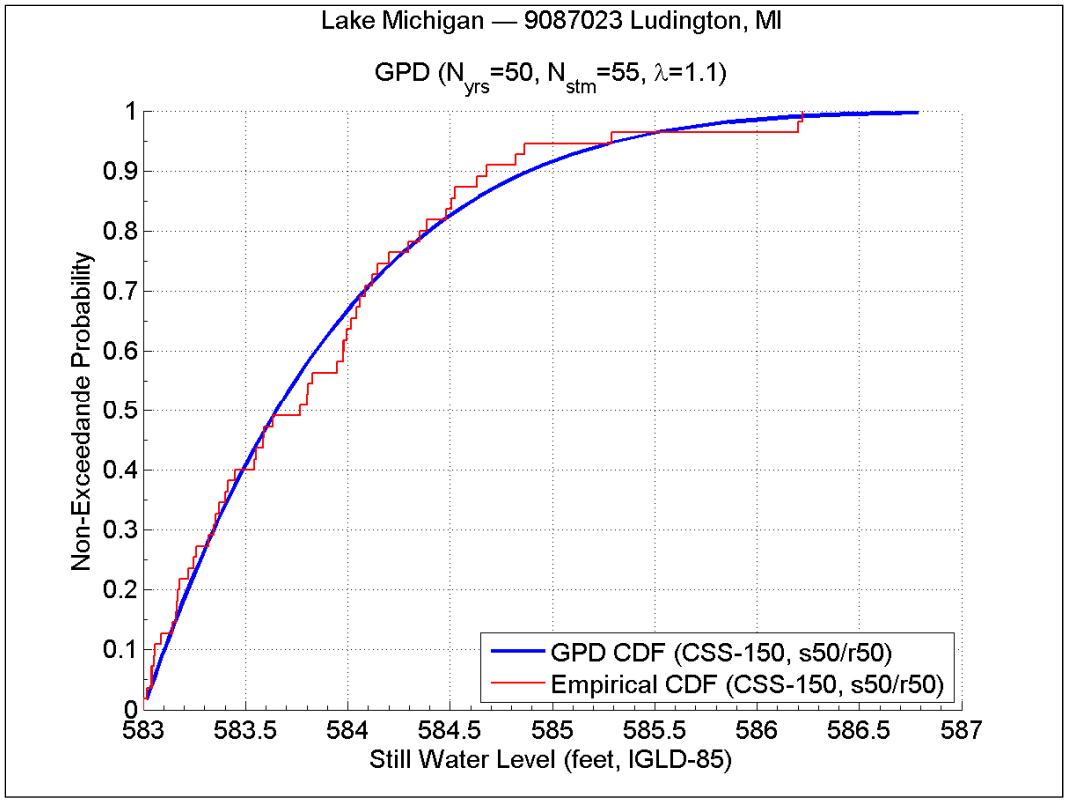



Figure B2. TWL CDF plot from CSS (50s/50r, with convective storms) for Ludington, MI.

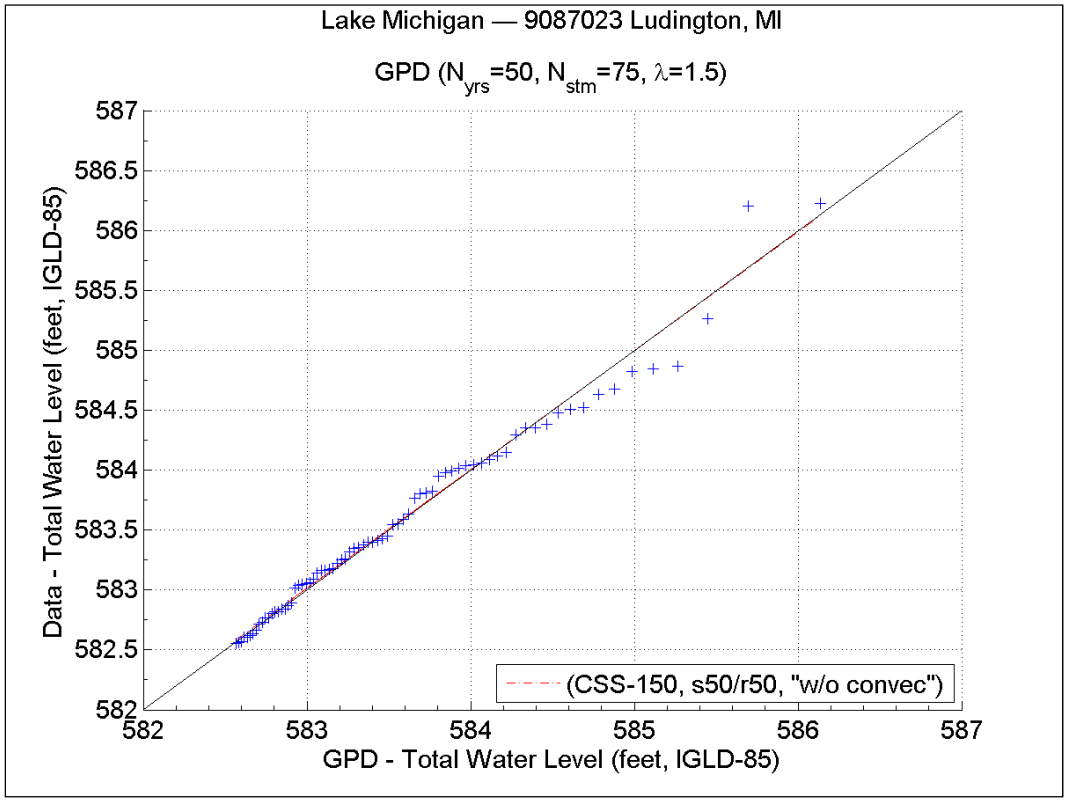



Figure B3. TWL Q-Q plot from CSS (50s/50r, no convective storms) for Ludington, MI.

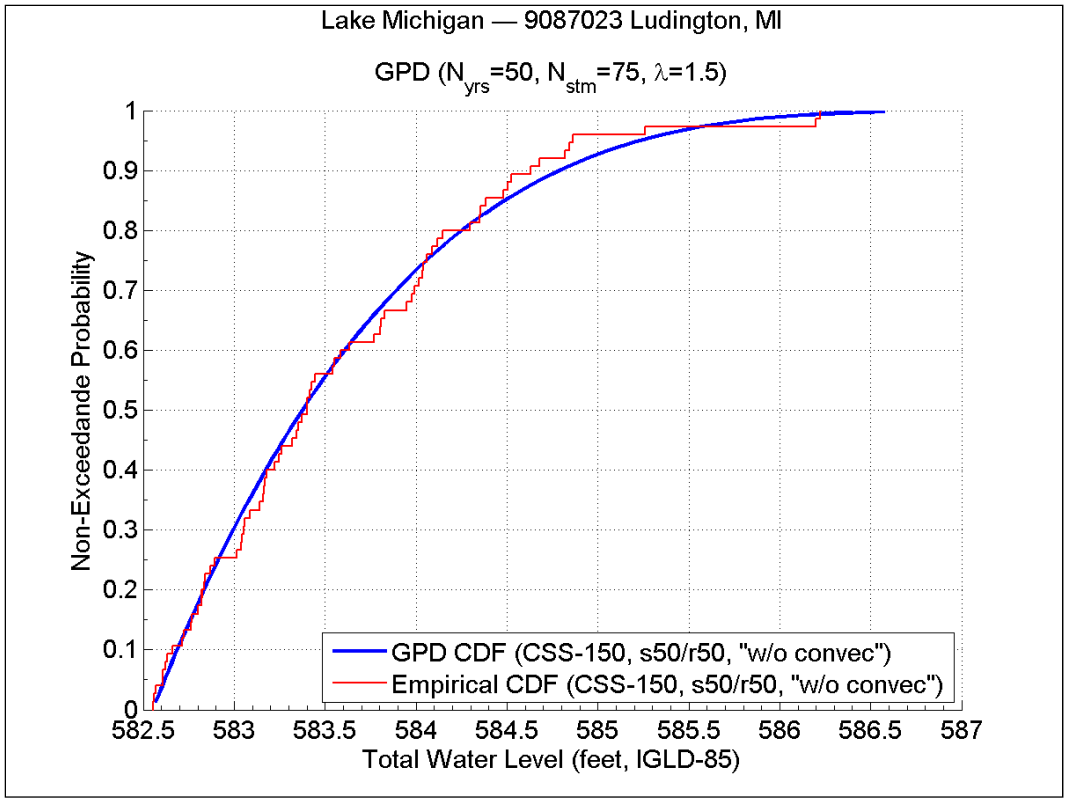



Figure B4. TWL CDF plot from CSS (50s/50r, no convective storms) for Ludington, MI.

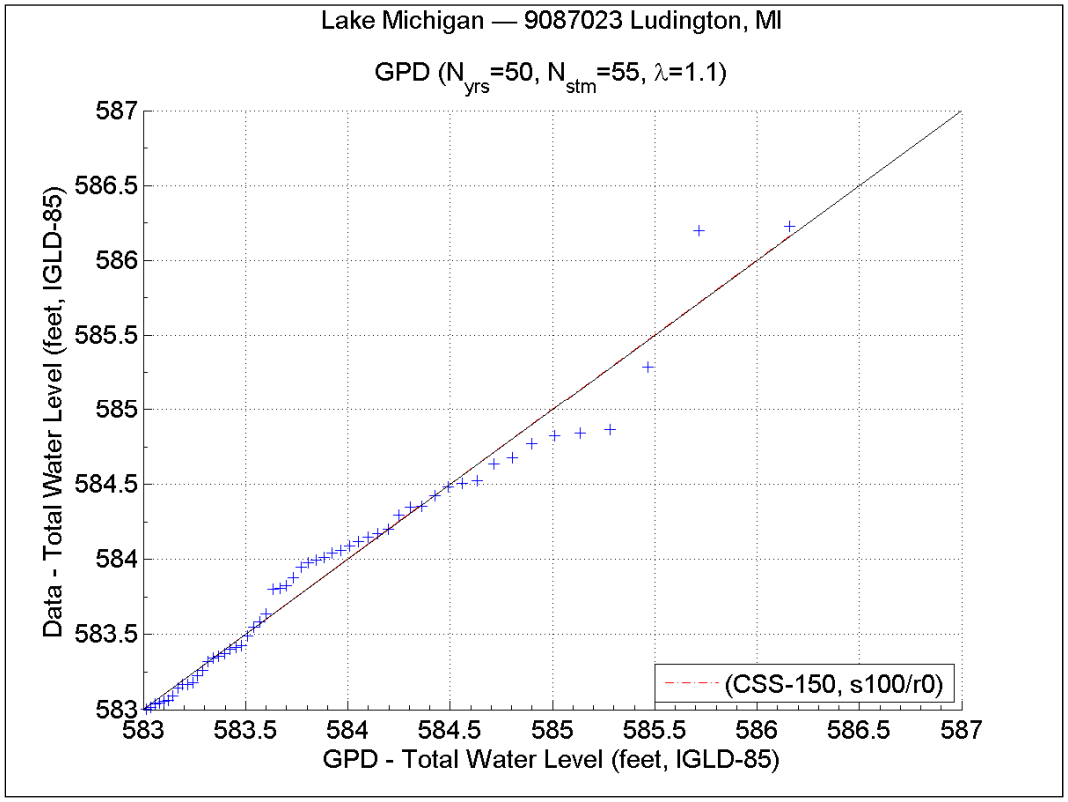



Figure B5. TWL Q-Q plot from CSS (100s/0r, with convective storms) for Ludington, MI.

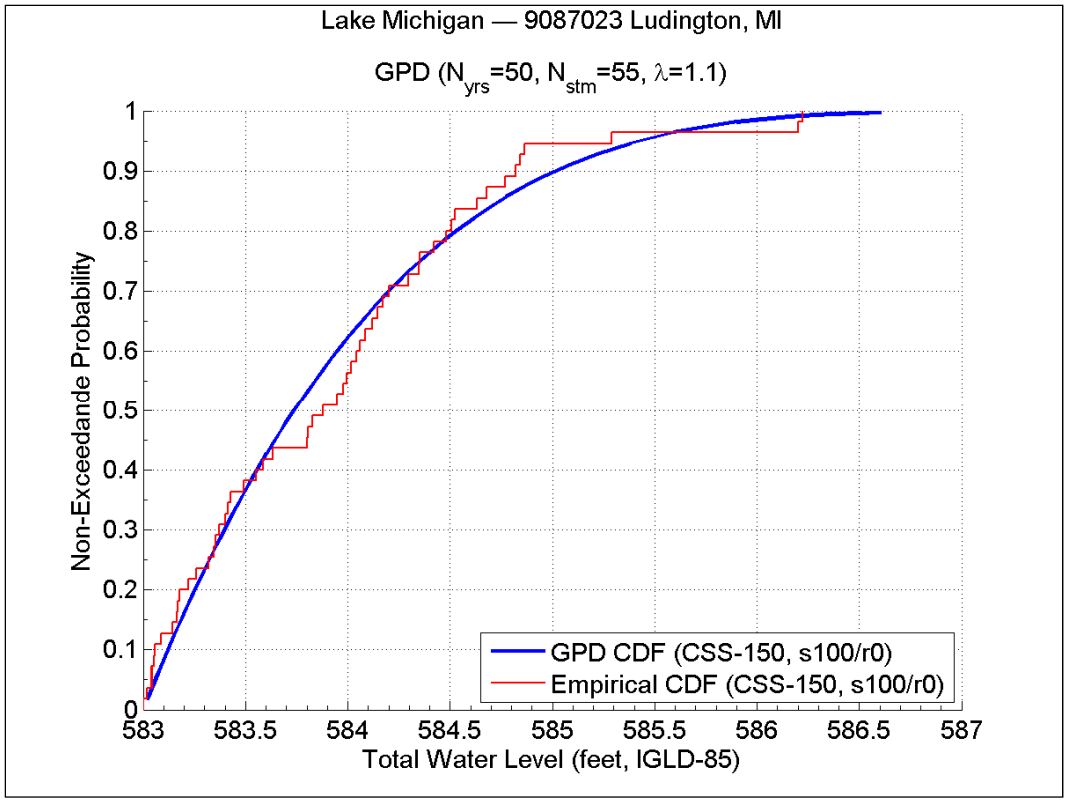



Figure B6. TWL CDF plot from CSS (100s/0r, with convective storms) for Ludington, MI.

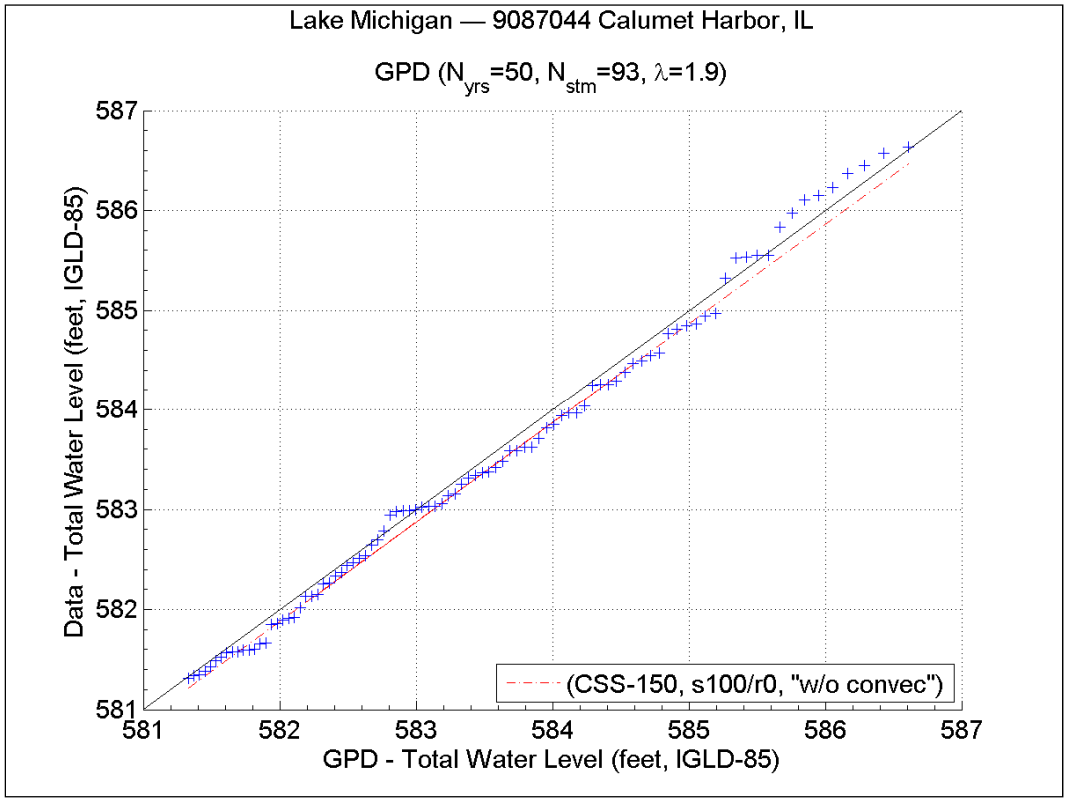



Figure B7. TWL Q-Q plot from CSS (100s/0r, no convective storms) for Ludington, MI.

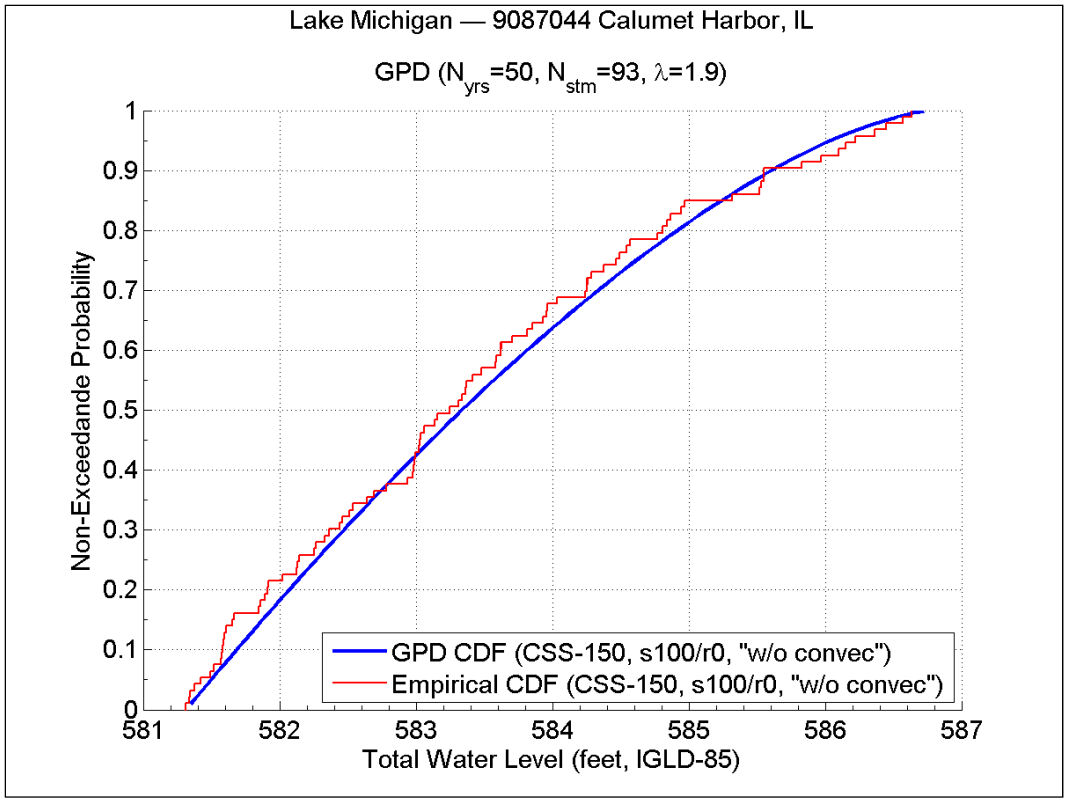



Figure B8. TWL CDF plot from CSS (100s/0r, no convective storms) for Ludington, MI.

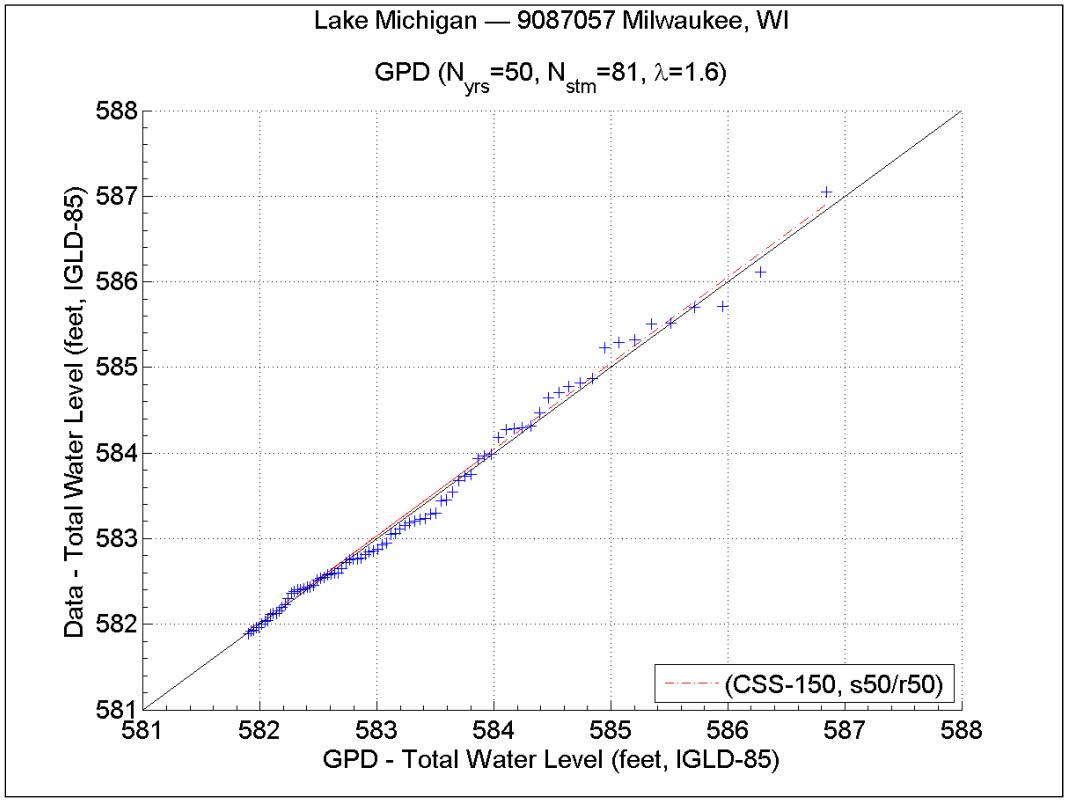



Figure B9. TWL Q-Q plot from CSS (50s/50r, with convective storms) for Milwaukee, WI.

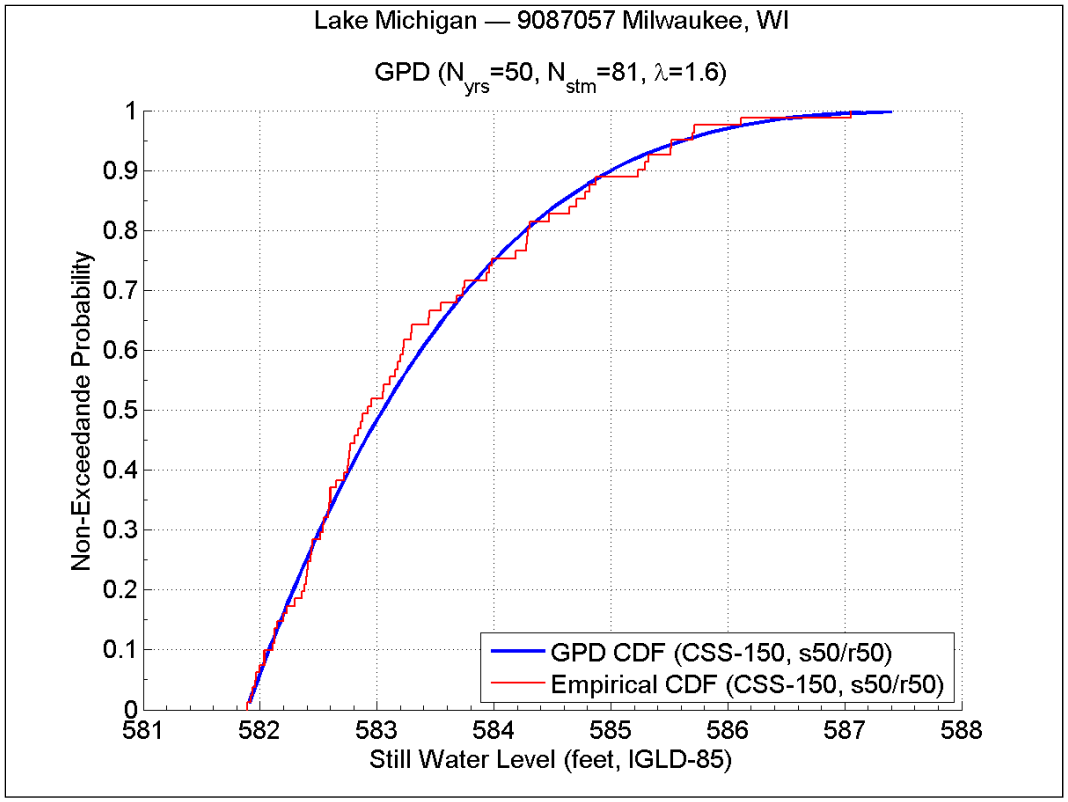



Figure B10. TWL CDF plot from CSS (50s/50r, with convective storms) for Milwaukee, WI.

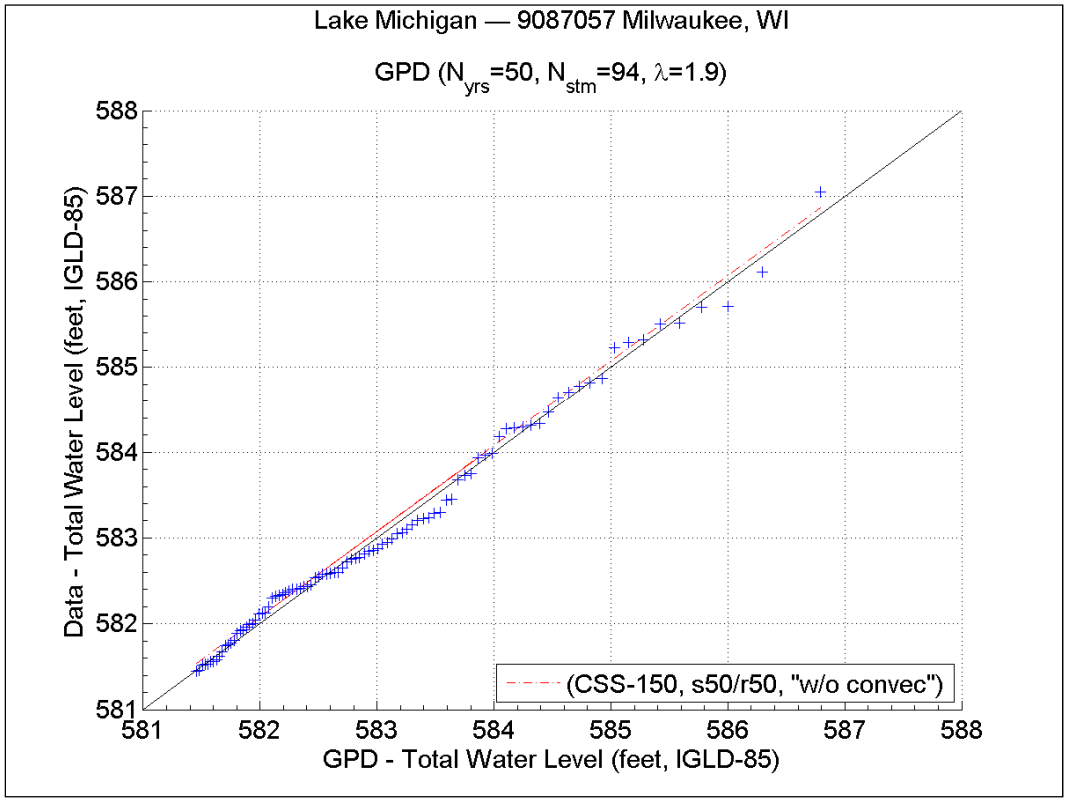



Figure B11. TWL Q-Q plot from CSS (50s/50r, no convective storms) for Milwaukee, WI.

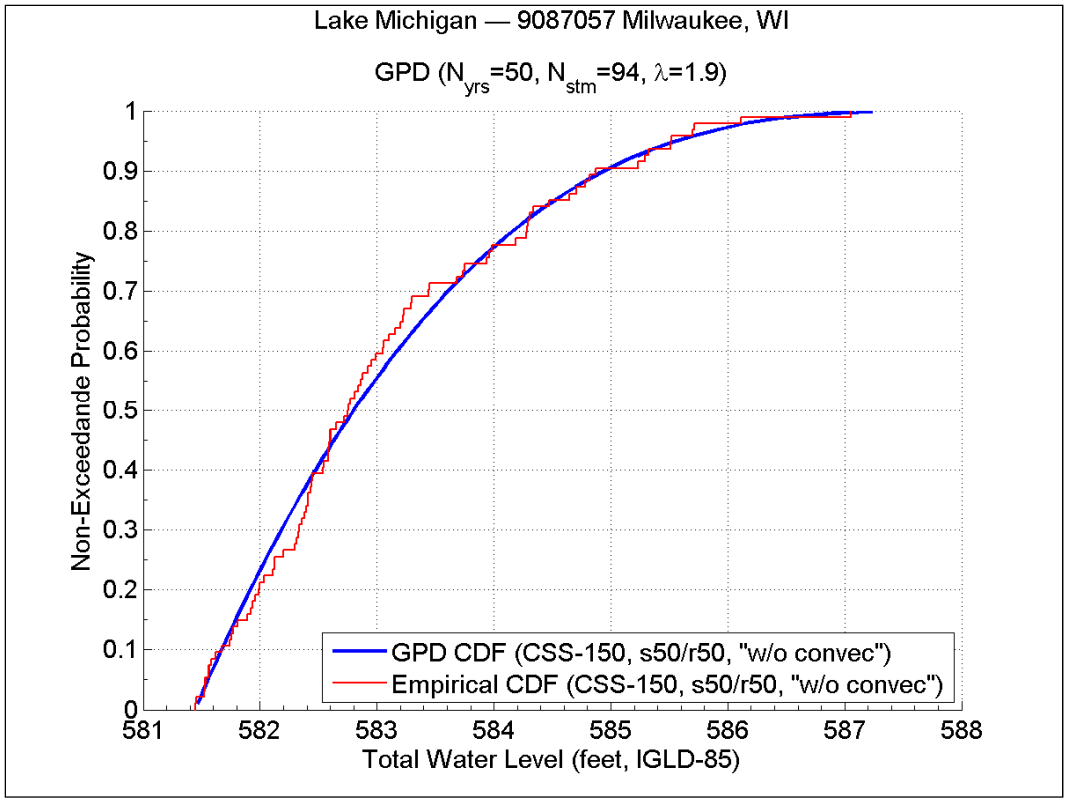



Figure B12. TWL CDF plot from CSS (50s/50r, no convective storms) for Milwaukee, WI.

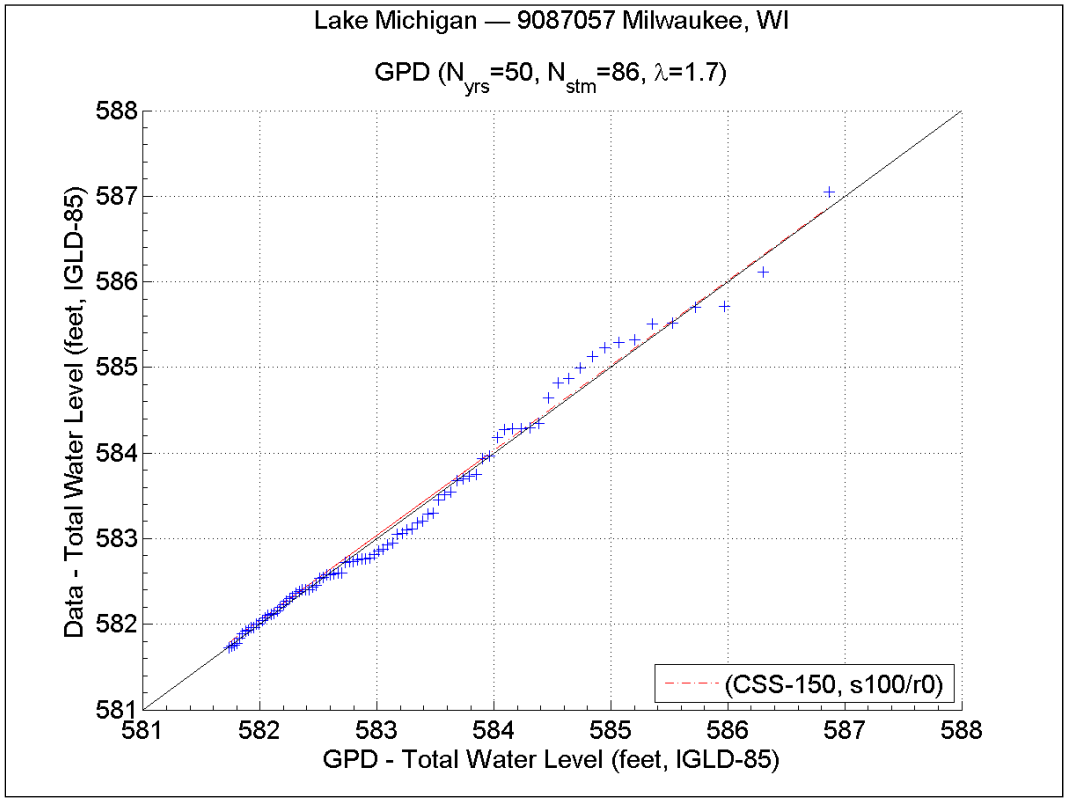



Figure B13. TWL Q-Q plot from CSS (100s/0r, with convective storms) for Milwaukee, WI.

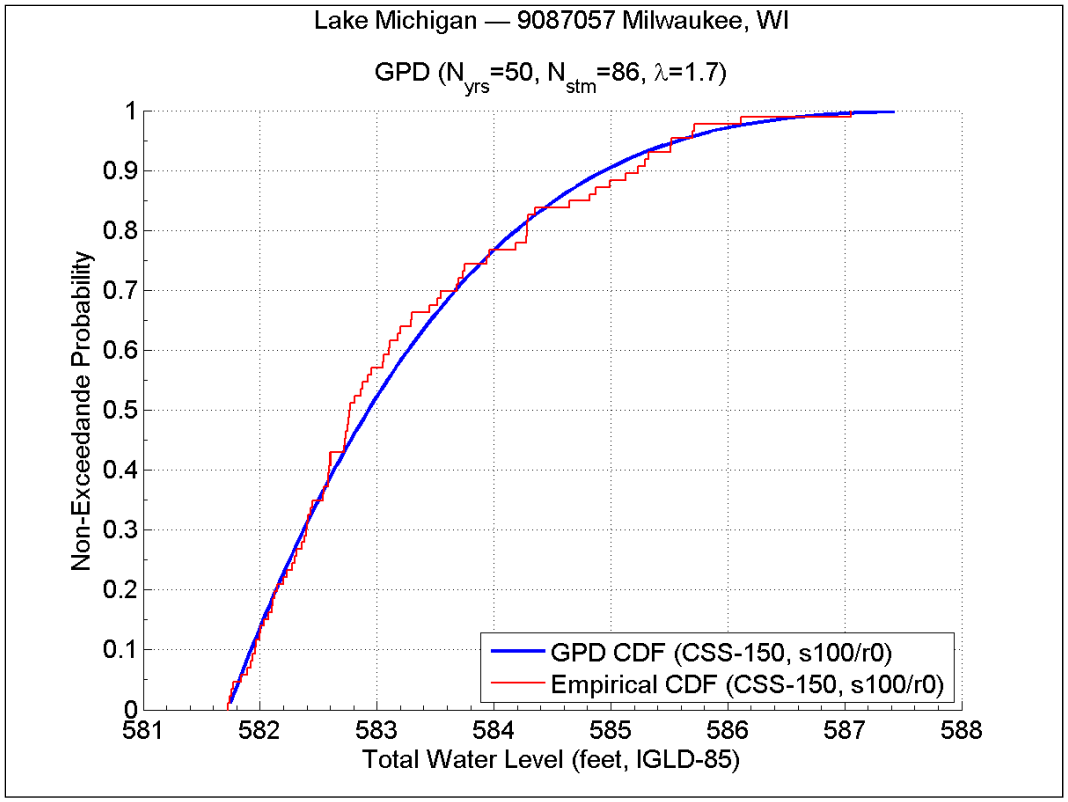



Figure B14. TWL CDF plot from CSS (100s/0r, with convective storms) for Milwaukee, WI.

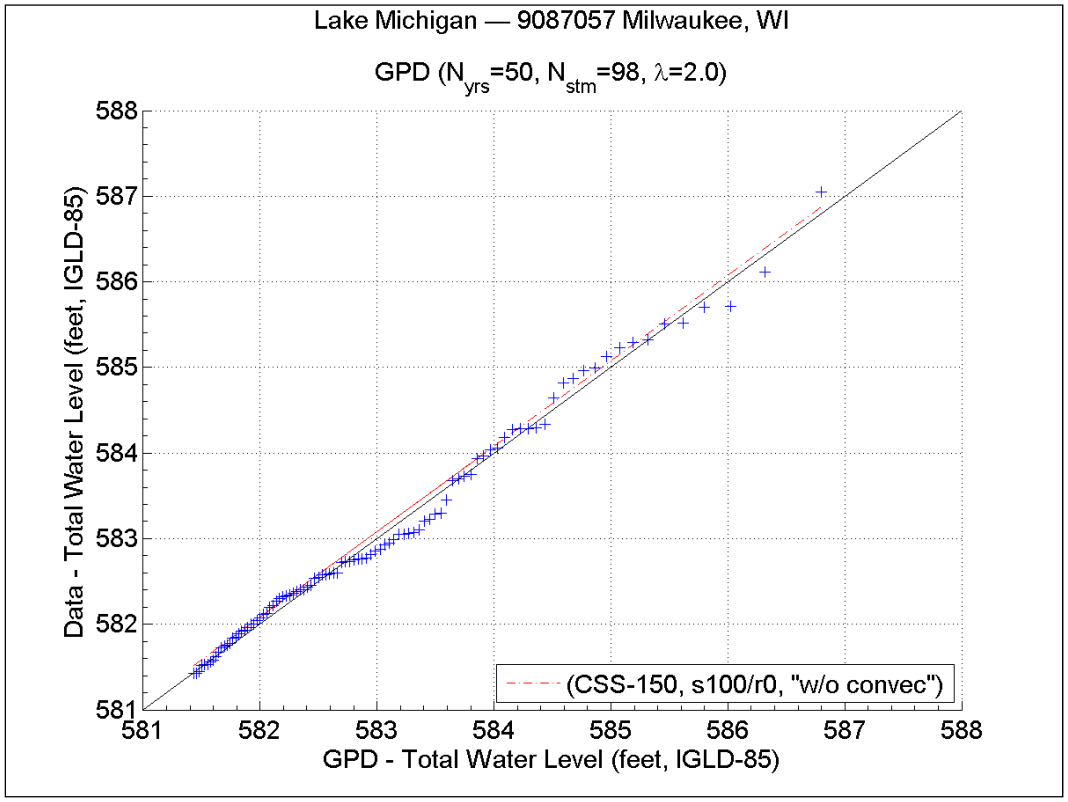



Figure B15. TWL Q-Q plot from CSS (100s/0r, no convective storms) for Milwaukee, WI.

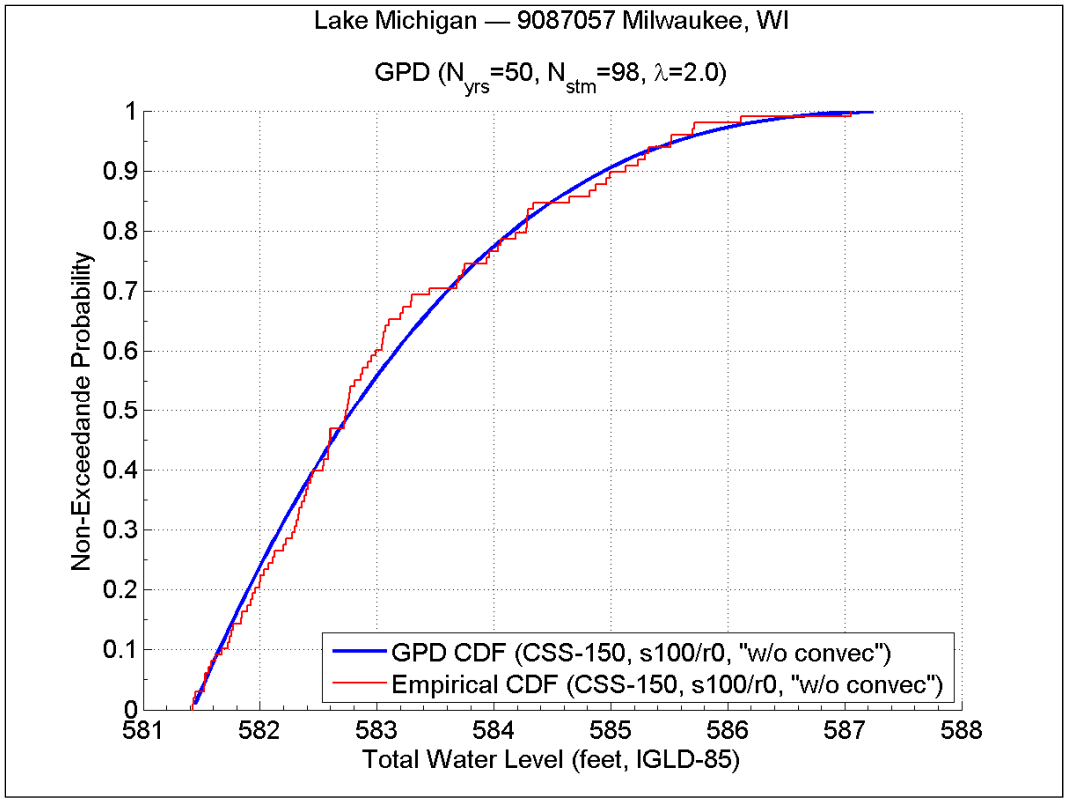



Figure B16. TWL CDF plot from CSS (100s/0r, no convective storms) for Milwaukee, WI.

# Appendix C: Lake Michigan 150-storm CSS

The following list (in chronological order) constitutes the final 150 storms selected as the CSS for Lake Michigan Flood Hazard Mapping production. Note that for *dual threat* events, dates correspond to either peak surge or peak wave, whichever occurred last.

| Storm # | Peak Date (UTC) | | | | Surge Ranking | Wave Ranking |
|---|---|---|---|---|---|---|
| | Year | Month | Day | Hour | | |
| 1 | 1960 | 1 | 15 | 16 | 5 | 19 |
| 2 | 1960 | 2 | 10 | 18 | 29 | 1 |
| 3 | 1960 | 3 | 22 | 7 | 2 | 10 |
| 4 | 1960 | 4 | 19 | 18 | 15 | 8 |
| 5 | 1961 | 4 | 20 | 19 | 6 | 8 |
| 6 | 1961 | 11 | 3 | 7 | 7 | 1 |
| 7 | 1962 | 1 | 14 | 9 | 1 | 12 |
| 8 | 1962 | 4 | 26 | 22 | n/a | 2 |
| 9 | 1962 | 5 | 23 | 21 | 28 | 10 |
| 10 | 1963 | 3 | 21 | 9 | 7 | 16 |
| 11 | 1963 | 4 | 4 | 0 | 7 | 7 |
| 12 | 1964 | 2 | 29 | 18 | 1 | 30 |
| 13 | 1964 | 3 | 5 | 9 | 18 | 17 |
| 14 | 1964 | 4 | 14 | 12 | n/a | 4 |
| 15 | 1964 | 5 | 6 | 22 | 8 | 10 |
| 16 | 1964 | 10 | 4 | 4 | 40 | 25 |
| 17 | 1965 | 2 | 25 | 21 | n/a | 9 |
| 18 | 1965 | 11 | 27 | 10 | 5 | 5 |
| 19 | 1965 | 12 | 25 | 15 | 1 | 9 |
| 20 | 1966 | 4 | 1 | 2 | 4 | n/a |
| 21 | 1966 | 11 | 29 | 6 | 7 | 6 |
| 22 | 1967 | 1 | 17 | 6 | 11 | 8 |
| 23 | 1967 | 1 | 27 | 3 | n/a | 10 |
| 24 | 1968 | 2 | 3 | 3 | n/a | 10 |
| 25 | 1968 | 4 | 7 | 22 | n/a | 4 |

| Storm # | Peak Date (UTC) | | | | Surge Ranking | Wave Ranking |
|---|---|---|---|---|---|---|
| | Year | Month | Day | Hour | | |
| 26 | 1968 | 12 | 13 | 18 | 7 | 5 |
| 27 | 1970 | 2 | 2 | 16 | 37 | 13 |
| 28 | 1970 | 2 | 16 | 22 | n/a | 8 |
| 29 | 1970 | 4 | 2 | 12 | n/a | 3 |
| 30 | 1970 | 10 | 10 | 4 | n/a | 28 |
| 31 | 1970 | 11 | 3 | 6 | 8 | n/a |
| 32 | 1970 | 11 | 25 | 22 | n/a | 28 |
| 33 | 1970 | 12 | 4 | 10 | 15 | 29 |
| 34 | 1971 | 1 | 26 | 15 | 6 | 2 |
| 35 | 1971 | 2 | 5 | 12 | 13 | 17 |
| 36 | 1971 | 2 | 27 | 15 | 28 | 3 |
| 37 | 1971 | 12 | 15 | 23 | 15 | n/a |
| 38 | 1971 | 12 | 30 | 20 | 3 | n/a |
| 39 | 1972 | 1 | 25 | 10 | 17 | 5 |
| 40 | 1972 | 1 | 25 | 12 | 17 | 12 |
| 41 | 1972 | 10 | 23 | 9 | 12 | 16 |
| 42 | 1972 | 12 | 31 | 0 | 14 | n/a |
| 43 | 1973 | 4 | 10 | 0 | 5 | 3 |
| 44 | 1973 | 10 | 15 | 19 | 5 | n/a |
| 45 | 1973 | 10 | 28 | 14 | 19 | n/a |
| 46 | 1973 | 12 | 13 | 21 | n/a | 6 |
| 47 | 1974 | 4 | 3 | 22 | n/a | 8 |
| 48 | 1975 | 1 | 12 | 0 | 1 | 1 |
| 49 | 1975 | 2 | 24 | 9 | n/a | 8 |
| 50 | 1975 | 11 | 10 | 18 | 2 | 21 |
| 51 | 1975 | 11 | 30 | 15 | 23 | 17 |
| 52 | 1976 | 1 | 28 | 10 | 31 | 6 |
| 53 | 1977 | 11 | 11 | 12 | n/a | 19 |
| 54 | 1978 | 1 | 9 | 21 | n/a | 9 |
| 55 | 1978 | 1 | 27 | 0 | n/a | 2 |
| 56 | 1978 | 5 | 13 | 15 | 50 | 27 |
| 57 | 1978 | 10 | 25 | 4 | n/a | 7 |

| Storm # | Peak Date (UTC) | | | | Surge Ranking | Wave Ranking |
|---|---|---|---|---|---|---|
| | Year | Month | Day | Hour | | |
| 58 | 1978 | 12 | 20 | 20 | 9 | n/a |
| 59 | 1979 | 4 | 12 | 8 | 10 | 43 |
| 60 | 1979 | 11 | 1 | 21 | 37 | 19 |
| 61 | 1979 | 12 | 25 | 6 | 20 | 2 |
| 62 | 1980 | 10 | 23 | 0 | 38 | 31 |
| 63 | 1981 | 5 | 11 | 7 | n/a | 8 |
| 64 | 1981 | 10 | 1 | 15 | 10 | 30 |
| 65 | 1982 | 1 | 4 | 19 | 5 | 14 |
| 66 | 1982 | 1 | 23 | 21 | 5 | 29 |
| 67 | 1982 | 4 | 4 | 6 | 4 | 45 |
| 68 | 1982 | 10 | 20 | 15 | 16 | n/a |
| 69 | 1983 | 1 | 15 | 9 | 25 | 43 |
| 70 | 1983 | 2 | 2 | 21 | n/a | 21 |
| 71 | 1983 | 4 | 1 | 0 | n/a | n/a |
| 72 | 1983 | 4 | 14 | 16 | n/a | 14 |
| 73 | 1983 | 5 | 7 | 22 | 18 | 3 |
| 74 | 1983 | 11 | 11 | 9 | 39 | 19 |
| 75 | 1983 | 11 | 28 | 22 | 11 | n/a |
| 76 | 1984 | 2 | 28 | 21 | 41 | 4 |
| 77 | 1984 | 4 | 16 | 13 | n/a | 14 |
| 78 | 1985 | 1 | 25 | 17 | 36 | n/a |
| 79 | 1985 | 3 | 5 | 4 | 3 | n/a |
| 80 | 1985 | 3 | 31 | 22 | 17 | n/a |
| 81 | 1985 | 11 | 20 | 7 | 1 | n/a |
| 82 | 1985 | 12 | 2 | 15 | 2 | 35 |
| 83 | 1987 | 2 | 8 | 21 | 9 | 4 |
| 84 | 1987 | 3 | 10 | 0 | 1 | 21 |
| 85 | 1987 | 12 | 15 | 15 | 2 | 7 |
| 86 | 1988 | 1 | 20 | 14 | 35 | 10 |
| 87 | 1988 | 4 | 6 | 21 | 40 | 17 |
| 88 | 1988 | 10 | 31 | 21 | n/a | 10 |
| 89 | 1988 | 11 | 17 | 3 | 4 | 2 |

| Storm # | Peak Date (UTC) | | | | Surge Ranking | Wave Ranking |
|---|---|---|---|---|---|---|
| | Year | Month | Day | Hour | | |
| 90 | 1989 | 1 | 22 | 3 | n/a | 2 |
| 91 | 1997 | 2 | 22 | 4 | 50 | 15 |
| 92 | 1989 | 11 | 16 | 18 | n/a | 2 |
| 93 | 1989 | 11 | 30 | 0 | 18 | 44 |
| 94 | 1989 | 12 | 26 | 4 | 8 | 11 |
| 95 | 1990 | 1 | 29 | 6 | 12 | 30 |
| 96 | 1990 | 2 | 27 | 0 | 15 | 5 |
| 97 | 1990 | 3 | 16 | 21 | n/a | 39 |
| 98 | 1990 | 4 | 10 | 19 | n/a | 8 |
| 99 | 1990 | 12 | 4 | 18 | 1 | 28 |
| 100 | 1991 | 1 | 23 | 6 | n/a | 5 |
| 101 | 1991 | 3 | 2 | 19 | n/a | 2 |
| 102 | 1991 | 11 | 2 | 9 | 13 | 14 |
| 103 | 1992 | 3 | 10 | 18 | 14 | 23 |
| 104 | 1992 | 5 | 23 | 22 | n/a | 2 |
| 105 | 1992 | 11 | 3 | 15 | 4 | 27 |
| 106 | 1992 | 12 | 25 | 15 | n/a | 4 |
| 107 | 1993 | 4 | 1 | 13 | 13 | 2 |
| 108 | 1993 | 4 | 20 | 12 | 4 | 37 |
| 109 | 1993 | 11 | 5 | 7 | 13 | n/a |
| 110 | 1993 | 11 | 19 | 18 | n/a | 2 |
| 111 | 1994 | 2 | 23 | 18 | n/a | 8 |
| 112 | 1994 | 4 | 17 | 0 | n/a | 14 |
| 113 | 1994 | 10 | 1 | 1 | 11 | n/a |
| 114 | 1994 | 11 | 18 | 21 | n/a | 14 |
| 115 | 2004 | 12 | 13 | 11 | n/a | 17 |
| 116 | 1994 | 11 | 28 | 18 | 7 | 3 |
| 117 | 1995 | 4 | 12 | 4 | 12 | n/a |
| 118 | 1995 | 11 | 11 | 15 | 22 | 39 |
| 119 | 1995 | 11 | 28 | 0 | 13 | 29 |
| 120 | 1996 | 1 | 29 | 22 | 21 | 36 |
| 121 | 1996 | 3 | 20 | 19 | 26 | 6 |

| Storm # | Peak Date (UTC) | | | | Surge Ranking | Wave Ranking |
|---|---|---|---|---|---|---|
| | Year | Month | Day | Hour | | |
| 122 | 1996 | 4 | 15 | 12 | 6 | n/a |
| 123 | 1996 | 10 | 31 | 0 | 16 | 28 |
| 124 | 1997 | 4 | 7 | 6 | n/a | 7 |
| 125 | 1997 | 5 | 1 | 1 | 19 | 15 |
| 126 | 1998 | 1 | 9 | 5 | 49 | n/a |
| 127 | 1998 | 3 | 9 | 16 | 5 | 7 |
| 128 | 1998 | 5 | 31 | 13 | 4 | n/a |
| 129 | 1998 | 11 | 11 | 0 | 3 | 1 |
| 130 | 1999 | 1 | 3 | 0 | 6 | 26 |
| 131 | 1999 | 2 | 12 | 7 | 23 | 6 |
| 132 | 2000 | 4 | 21 | 4 | 9 | 2 |
| 133 | 2001 | 4 | 12 | 4 | 16 | n/a |
| 134 | 2001 | 10 | 26 | 1 | 15 | 28 |
| 135 | 2002 | 2 | 12 | 3 | 31 | n/a |
| 136 | 2002 | 2 | 20 | 22 | 31 | 6 |
| 137 | 2002 | 3 | 10 | 7 | 13 | 10 |
| 138 | 2002 | 5 | 9 | 16 | n/a | 4 |
| 139 | 2003 | 11 | 4 | 20 | 28 | n/a |
| 140 | 2003 | 11 | 23 | 22 | 13 | n/a |
| 141 | 2005 | 11 | 16 | 8 | 4 | 11 |
| 142 | 2006 | 5 | 12 | 0 | 50 | n/a |
| 143 | 2006 | 11 | 16 | 13 | 34 | 10 |
| 144 | 2007 | 3 | 2 | 3 | 15 | n/a |
| 145 | 2007 | 4 | 12 | 2 | 16 | 12 |
| 146 | 2007 | 11 | 27 | 15 | 29 | 4 |
| 147 | 2007 | 12 | 23 | 16 | 6 | n/a |
| 148 | 2008 | 1 | 30 | 10 | n/a | 3 |
| 149 | 2009 | 4 | 6 | 16 | n/a | 1 |
| 150 | 2009 | 12 | 9 | 15 | 2 | 1 |

# Appendix D: Lake St. Clair – Statistical Analysis and Storm Sampling Approach

This Appendix summarizes the statistical analysis and sampling approach followed for Lake St. Clair and outlines how the implementation of this approach deviated from the methodology followed for Lake Michigan. Some of these differences result from: (1) the size of Lake St. Clair, and (2) a relatively limited set of data, when compared to the rest of the Great Lakes.

The surface area of Lake Michigan is roughly 52 times larger than Lake St. Clair (58,000 km2 vs. 1,114 km2). This is reflected in the availability of measured data. Whereas there are nine long-term NOAA-NOS water level gages in Lake Michigan, there are only two long-term gages ) on Lake St. Clair (i.e., St. Clair Shores, MI and Windmill Point, MI.

There are two additional long-term gages in the vicinity of Lake St. Clair (i.e., Algonac, MI and Fort Wayne, MI). However, these two gages are located within the boundaries of rivers adjacent to the lake. The Algonac gage is located within St. Clair River while the Fort Wayne gage is located in Detroit River.

Also, it is important to note there are no long-term wave data, neither measured or hindcast, available for Lake St. Clair.

## Data summary

### Datum

All Lake St. Clair water levels are referenced to the most recent datum, the International Great Lakes Datum of 1985 (IGLD-85). A description of the datum can be found on the USACE Detroit District web site[1].

### Water level data

Water level data were acquired from NOAA-NOS for Lake St. Clair for four stations. The coordinates of each of these stations are shown in Table D1.

[1] http://www.lre.usace.army.mil/greatlakes/hh/newsandinformation/igldatum1985/

Table D1. Coordinates of Water NOAA Water Level Stations for Lake St. Clair

| Station | Station Number | Latitude | Longitude |
|---|---|---|---|
| Algonac, MI | 9014070 | 42° 37.2' N | 82° 31.6' W |
| St. Clair Shores, MI | 9034052 | 42° 28.3' N | 82° 52.3' W |
| Fort Wayne, MI | 9044036 | 42° 17.9' N | 83° 50.5' W |
| Windmill Point, MI | 9044049 | 42° 21.4' N | 82° 55.8' W |

Long-term data were available for only four of these stations, as summarized in Tables D2 and D3. Data collected since 1970 are summarized in Table D2 and were collected at 6-minute or hourly intervals. Pre-1970 data are summarized in Table D3 and are available only as monthly averages and monthly maxima. The water level station locations are shown in Figure 1.

Table D2. Recording Periods for Measured Water Level Data for Lake St. Clair

| Station | Station Number | 6-Minute Records | Hourly Records |
|---|---|---|---|
| Algonac, MI | 9014070 | 1996-2010 | 1975-2010 |
| St. Clair Shores, MI | 9034052 | 1996-2010 | 1970-2010 |
| Fort Wayne, MI | 9044036 | 1996-2010 | 1970-2010 |
| Windmill Point, MI | 9044049 | 1999-2010 | 1970-2010 |

Table D3. Recording Periods for Monthly Measured Water Level Data for Lake St. Clair

| Station | Station Number | Record Monthly Average | Record Monthly Max. |
|---|---|---|---|
| Algonac, MI | 9014070 | 1901,1926-1929,1947, 1952-1969, 1975-2010 | 1952-1969, 1975-2010 |
| St. Clair Shores, MI | 9034052 | 1968-2010 | 1968-2010 |
| Fort Wayne, MI | 9044036 | 1970-2010 | 1970-2010 |
| Windmill Point, MI | 9044049 | 1897-2010 | 1952-2010 |

### Meteorological Data

Meteorological data were acquired from the NOAA National Climatic Data Center (NCDC). Only four U.S. land-based meteorological stations, and one Canadian station, covered the necessary record length, as summarized in Table D4.

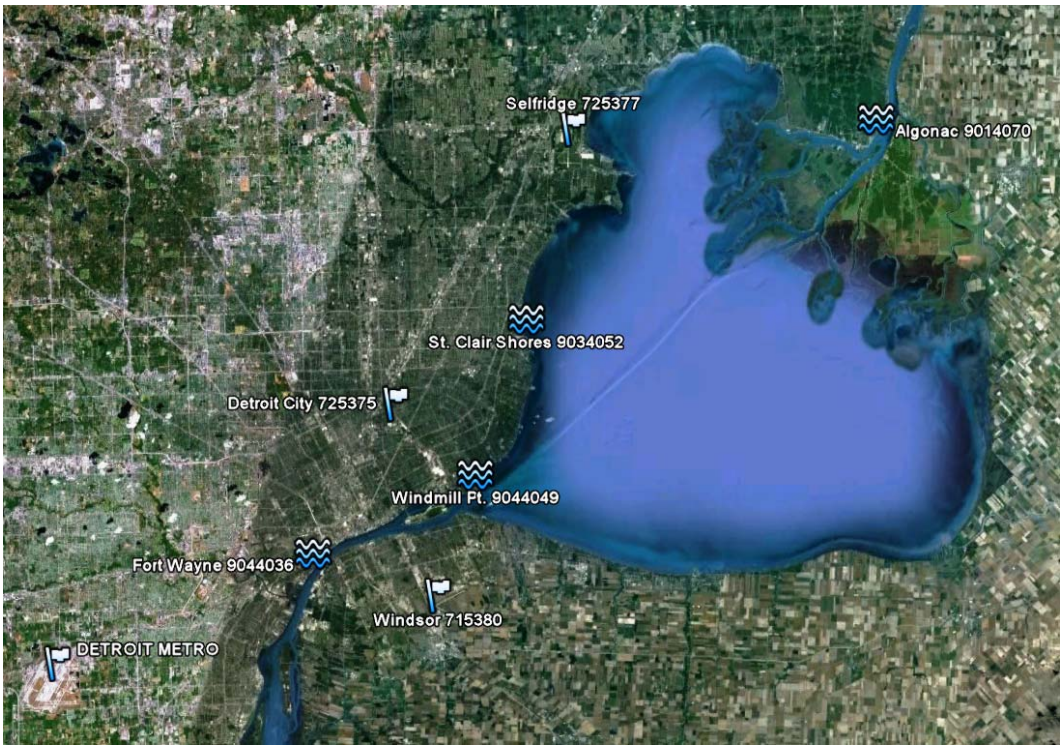



Figure D1. Location of NOAA water level gages and meteorological (met) stations on Lake St. Clair. Water level gages are denoted by three wavy lines, and met stations are waving flags.

Table D4. Recording Periods for Measured Meteorological Data for Lake St. Clair

| Station | Station Number | Latitude | Longitude | Hourly Records |
|---|---|---|---|---|
| Detroit Metro | 725370/94847 | 42.215˚ N | 83.349˚ W | 1937 – 2009 |
| Detroit City, MI | 725375/14822 | 42.409˚ N | 83.010˚ W | 1930 – 1948, 1966 – 2009 |
| Selfridge ANGB, MI | 725377/14804 | 42.613˚ N | 82.832˚ W | 1960 – 2009 |
| Windsor, CAN | 715380 | 42.267˚ N | 82.967˚ W | 1955 – 1963, 1973 – 2009 |

### Storm surrogates

Measured winds from the four long-term meteorological stations shown in Table D4 were used to estimate storm surge and wave heights with relatively simpler surrogate calculation methods and were then compared against their measured/more rigorously hindcast counterparts, following the same methodology developed for Lake Michigan.

The storm surrogate analysis performed for Lake Michigan demonstrated that:

1. The steady-state storm surge methodology, as discussed in Chapter 5, is not accurate enough to be used as part of the storm selection process. Nonetheless, the methodology was still applied for Lake St. Clair to reevaluate this conclusion.
2. Given the right conditions, the ACES/CEM wind-wave generation methodology (Chapter 6) could be used to estimate waves with a reasonable amount of certainty, in the absence of measured or hindcast wave data. However, this methodology proved to be inadequate for Lake St. Clair, for reasons to be discussed in the *Wind-wave results* Section.

#### Storm surge results

Similar to the case of Lake Michigan, although significant effort was put behind the calibration process, the steady-state surrogate storm surge approach did not have sufficient skill for storm screening/sampling. The dynamic effects involved in the estimation of storm surge are significant and not considering them leads to these inaccuracies. The use of numerical models that take into account the dynamic effects of storm surge would be required to obtain adequate estimates.

#### Wind-wave results

The main difficulty when applying the ACES/CEM wind-wave generation methodology was the lack of long-term measured waves to validate the surrogate wave estimates. A Canadian buoy (NOMAD 45174; 42.430 N 82.680 W), with 11-year record length (2000 – 2010), was used to try to generate some measure of skill for evaluating the surrogate estimates.

Figures D2 and D3 show the comparison between the Canadian buoy measurements and the wave surrogate estimates for the top 20 and top 100 storms, respectively, for the 2001-2010 period. From these plots it can be seen that the scatter is quite significant and, in some cases, there are differences of more that 100 percent between estimated and measured wave heights. From these results, and given the limited data available for calibration and validation purposes, the ACES/CEM wind-wave generation methodology cannot be recommended for Lake St. Clair.

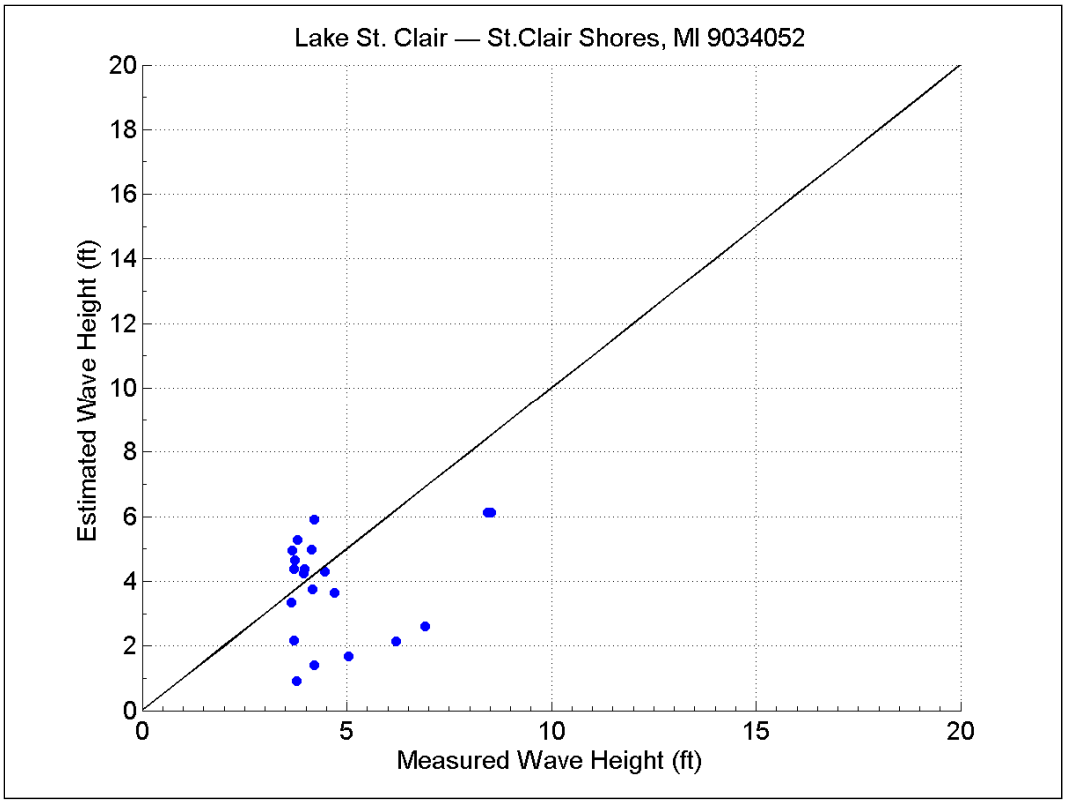



Figure D2. Top 20 wave events for St. Clair Shores, MI gage.

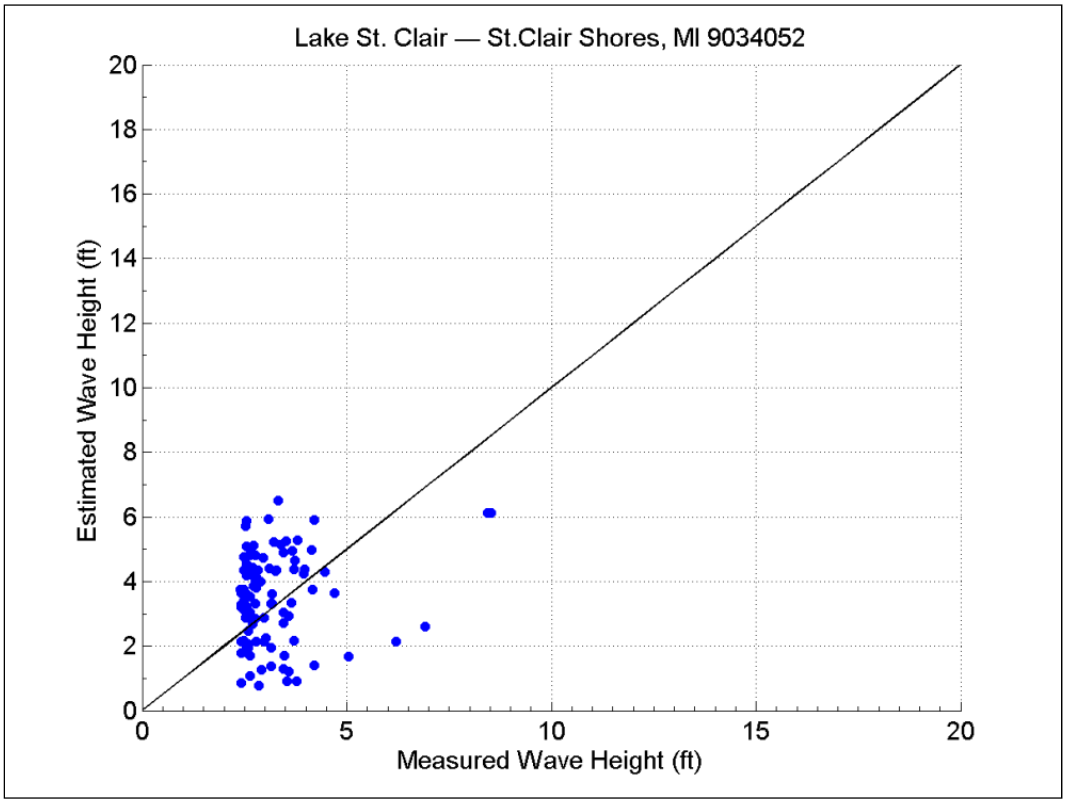



Figure D3. Top 100 wave events for St. Clair Shores, MI gage.

## Storm selection for flood mapping

The sampling approach used to select storms for Lake St. Clair follows as much as possible the methodology established for Lake Michigan, albeit some clear limitations.

Since neither long-term wave measurements nor storm surrogate estimates are available for Lake St. Clair, a surge-only POT approach was implemented. In other words, storms were sampled using a 100 percent surge/0 percent waves prioritization ratio. The adequacy of using this fractional sampling was broadly discussed in Chapters 12 and 13.

Storms were sampled from the only two long-term NOAA-NOS water level gages located within Lake St. Clair boundaries (i.e., stations 9034052 and 9044049). The final storm selection process that was modified and adopted for Lake St. Clair is discussed next.

### Initial identification of composite storm set

1. Identify storms having the highest peak storm surge at each of the long-term water level measurement sites using monthly maxima (highest hourly each month) and monthly average, and/or hourly water level data. Rank storms based on magnitude of peak surge.
2. Starting with highest 75 ranked surge events at each location, define highest ranked surge events for the lake. If storms are *duplicates* (i.e. a storm is a large event at one or more sites), reject duplicate event (or events) with the lowest site-specific rank and include the next largest storm surge event at that site. In general, the replacement procedure might involve going deeper in the rankings and sampling storms below the top 75.
3. Balance number of storms selected for each site to maximize consistency in geographical coverage of selected storms; define an initial set of 150 storms (roughly 50 percent of storms sampled from station 9034052 and 50 percent from 9044049).
4. Using the ice maps nearest the time of each of the 150 storms, or the ice maps that are bracketing the storm, record if shore-fast ice is present in each of the various coastline segments of the lake , and either partially or fully blocking waves in that segment (greater than 70 percent concentration assumed to be blocking waves).
5. Examine the peak surge and wind conditions (such as extreme winds and pressure differentials favorable to surge generation) during all these events before finalizing the storm set.

Note that after several cycles of ice-screening and replacement it was decided to generate a 144-storm CSS for Lake St. Clair.

## Sampling across high and low lake levels

The distribution of storms contained in the CSS across all lake levels for the 1960-2010 period are shown for the St. Clair Shores, MI and Windmill Point, MI gages in Figures D4 and D5, respectively. Also illustrated in the plots are the 1960-2010 monthly maxima water levels (red lines) and the NNR fits (blue lines).

As is the case for Lake Michigan, both figures show how well all three high lake-level periods (i.e., mid 1970s, mid 1980s, and late 1990s) and all low lake-level periods (mid 1960s, late 1970s, early 1990s and 2000s) are represented by the sampled storms. It can be seen that the CSS not only captures the decadal variation in water levels but also much of the higher frequency variation in water levels.

The highest overall water level for the St. Clair Shores station occurred during the high lake-level period of the mid 1980s. More specifically, this highest overall lake level occurred in October, 1986. However, no storm corresponding to that month and year was sampled through the POT approach, which means that if a storm occurred within that period, the magnitude of the surge it generated was below the established threshold.

Nonetheless, for statistical purposes, the highest overall water level must be accounted for. A manual revision of all storms revealed that a storm did occur on October 6, 1980. Although the magnitude of its surge was roughly 0.10 meters below threshold, it coincided with the highest lake level on record, resulting in the highest overall water level. So the October 6, 1980 storm was included as part of the final CSS.

## Still water level – full storm set

The Full Storm Set (FSS), as discussed in Chapter 10, is defined as a set of storms where all events are sampled from the same spatial location. The SWL-FSS are generated using the entire record length available at each gage location. Thus, for the St. Clair Shores gage, used in this section as an example, the record length is limited to the 42-year period between 1968 and 2009. For the Windmill Point gage) the FSS comprise the 58-year period between 1952 and 2009.

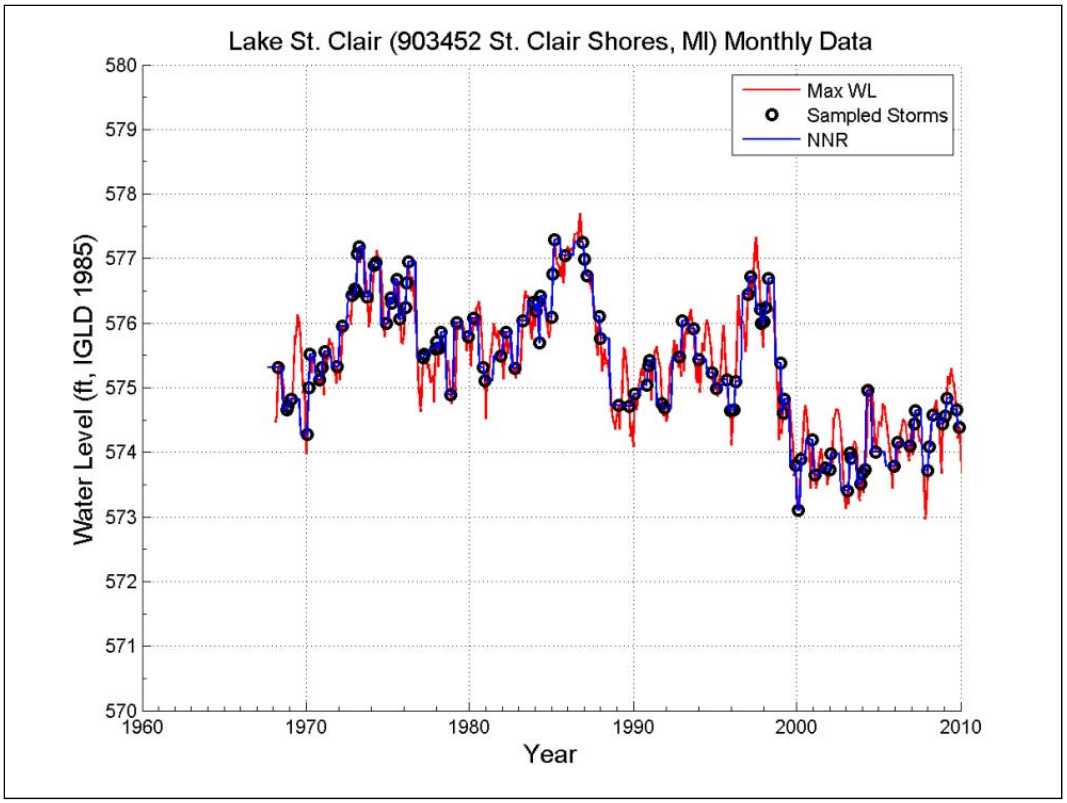



Figure D4. Distribution of 150-storm CSS across lake levels for St. Clair Shores, MI.

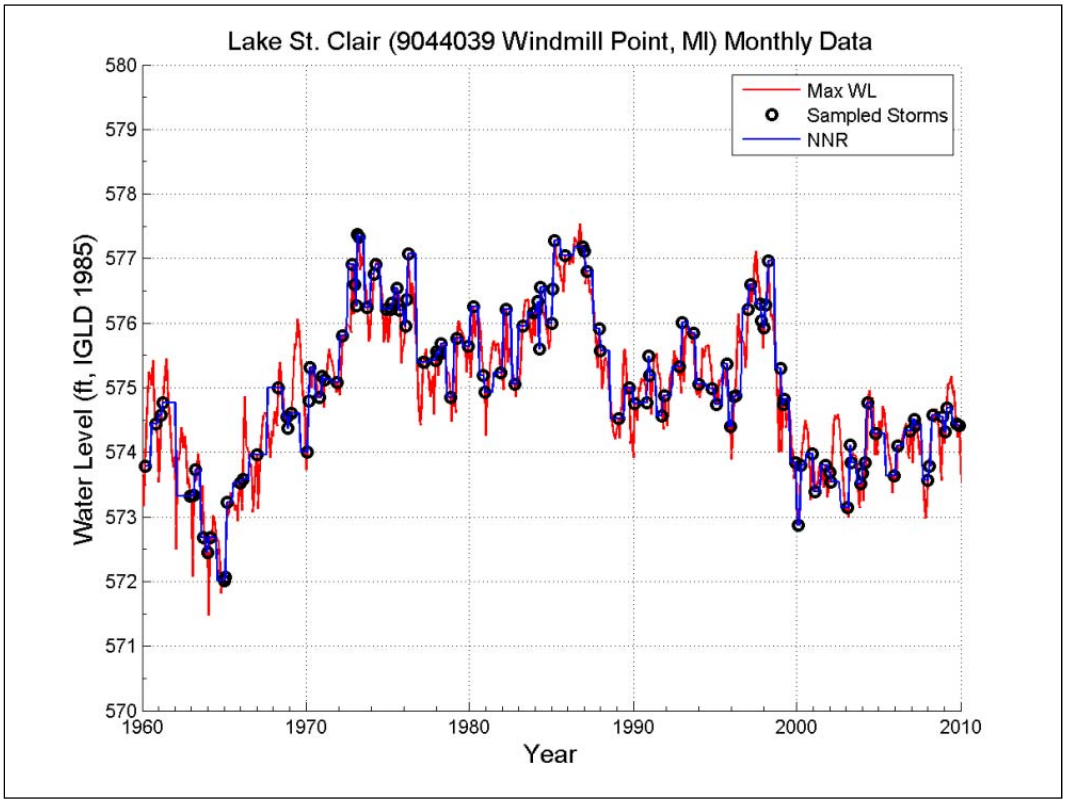



Figure D5. Distribution of 150-storm CSS across lake levels for Windmill Point, MI.

Figures D6 through D9 show the SWL-FSS Q-Q plots, CDF plots, and RP plots corresponding to St. Clair Shores water level gage. Likewise, Figures D9 through D11 show the plots for the Windmill Point gage.

## Still water level – composite storm set

The Composite Storm Set (CSS) is defined as a single set, or composite, of storms derived from storms sampled from different gage locations around a given lake. For Lake Michigan, it was determined that the necessary number of storms to generate the CSS was roughly 150, as discussed in Chapter 7 (*Storm Sample Size*). However, further analyses for Lake St. Clair showed that a slightly reduced set could be used instead. Therefore a set of 144 storms was chosen for Lake St. Clair. In this case, the reduced CSS is justified because: (1) the magnitudes of the storm surge observed in Lake St. Clair are relatively small, and (2) the differences between BFEs computed from 150-CSS and reduced CSS were negligible.

The gages of St. Clair Shores and Windmill Point observed maximum surge elevations of just 0.97 ft and 1.45 ft, respectively. Also, the magnitudes of storm surges due the high-frequency events showed little

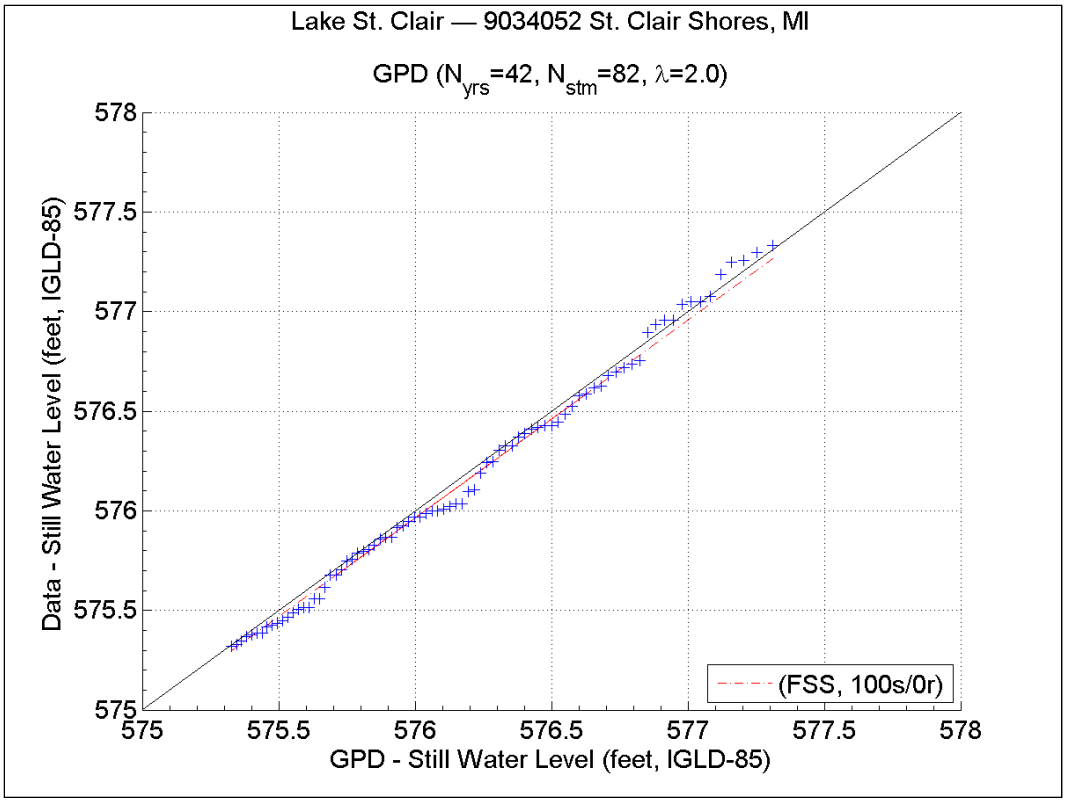



Figure D6. SWL Q-Q plot from FSS (with convective storms) for St. Clair Shores, MI.

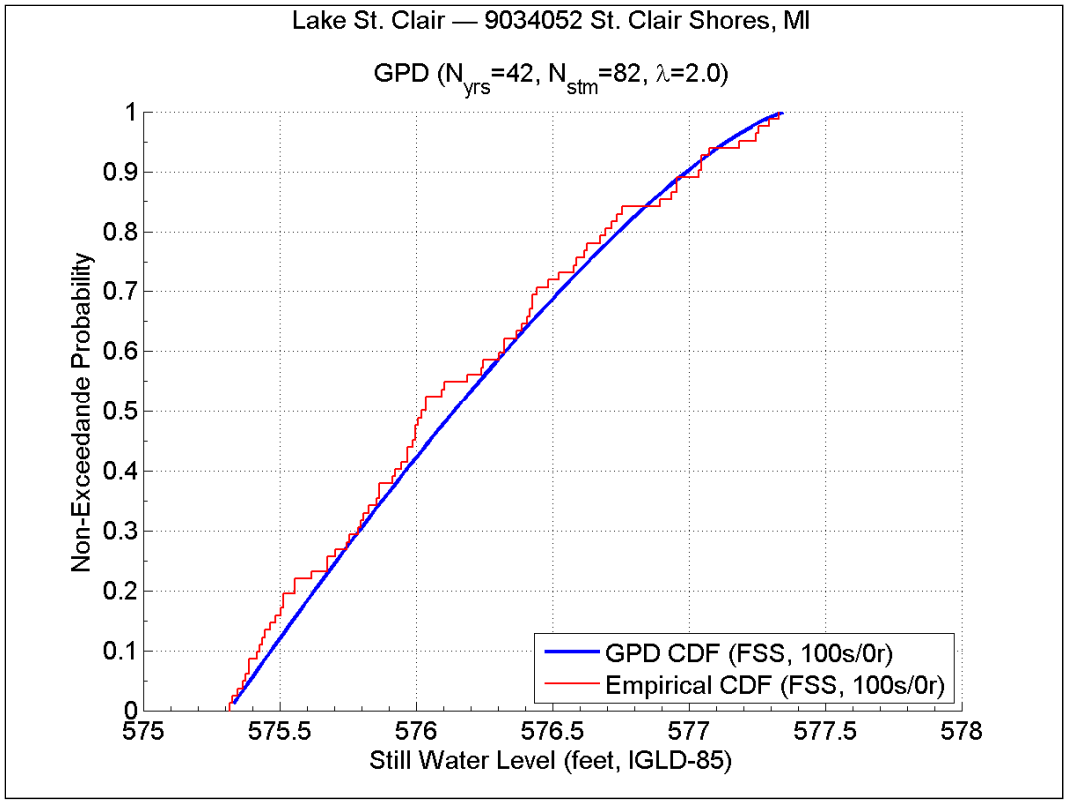



Figure D7. SWL CDF plot from FSS (with convective storms) for St. Clair Shores, MI.

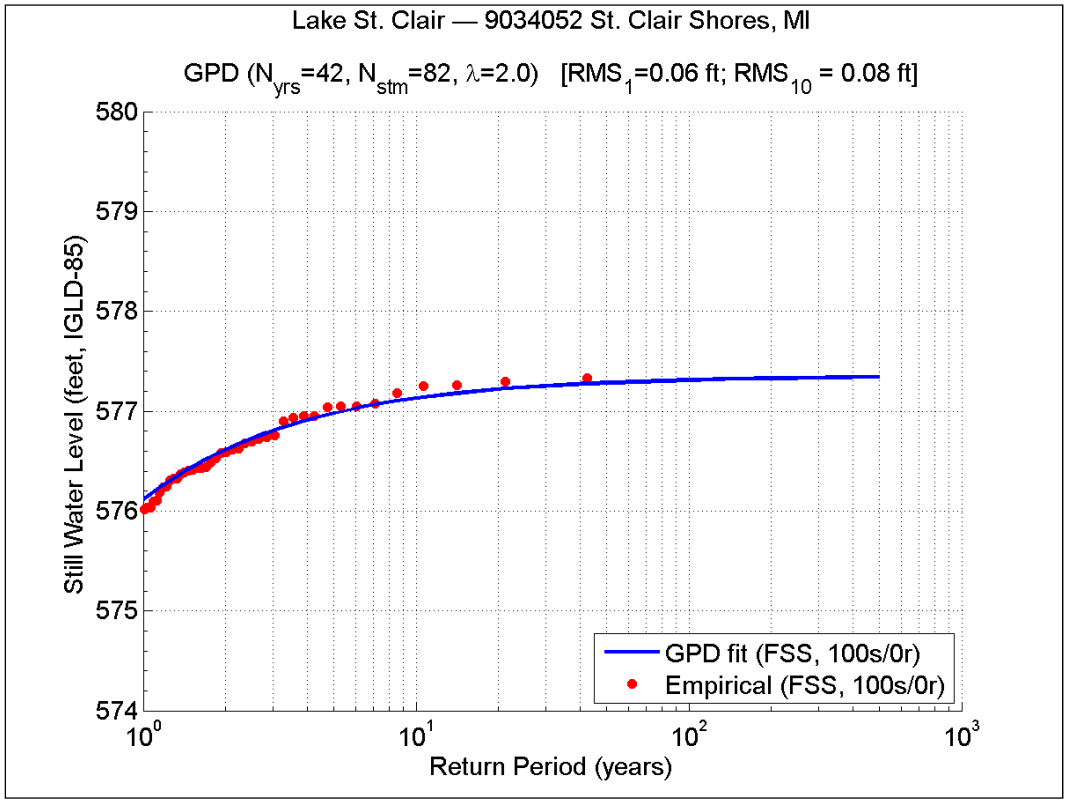



Figure D8. SWL RP plot from FSS (with convective storms) for St. Clair Shores, MI.

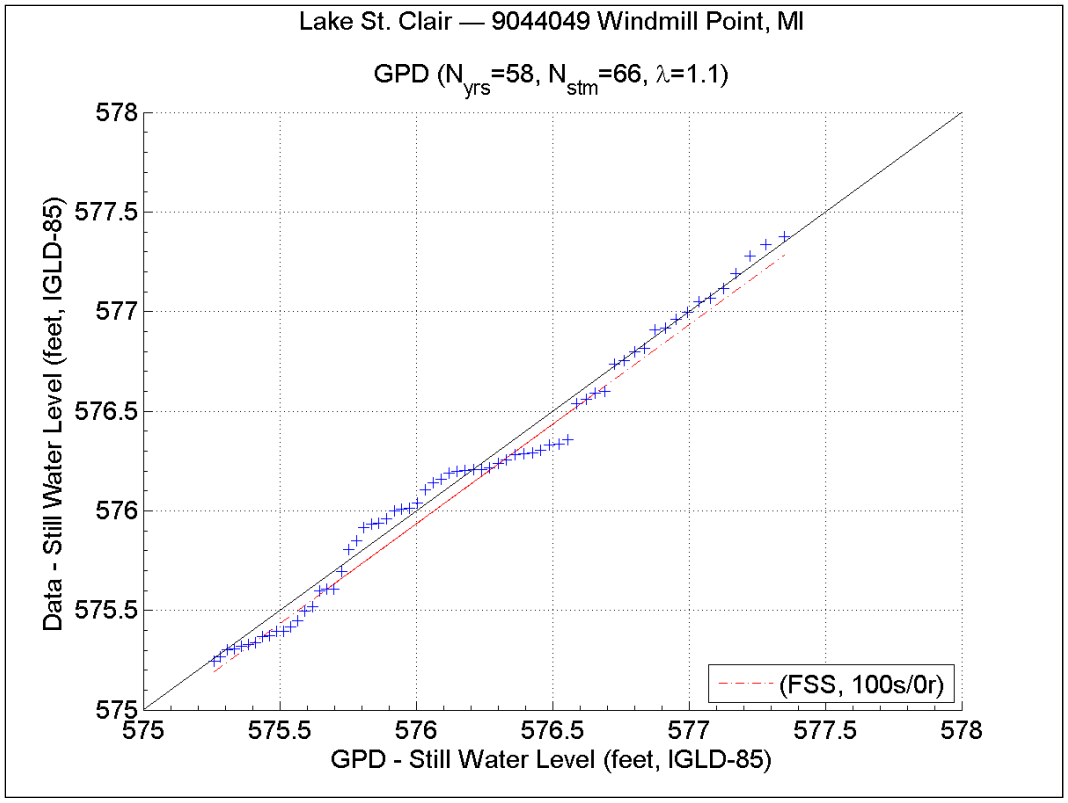



Figure D9. SWL Q-Q plot from FSS (with convective storms) for Windmill Point, MI.

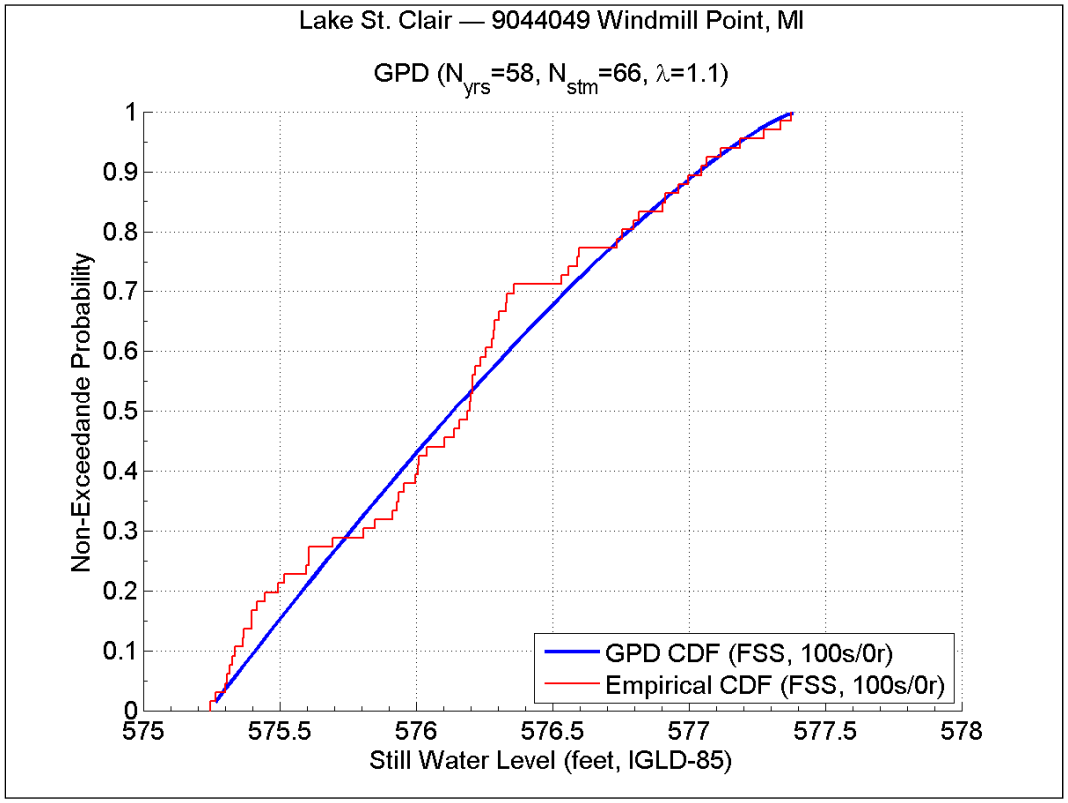



Figure D10. SWL CFD plot from FSS (with convective storms) for Windmill Point, MI.

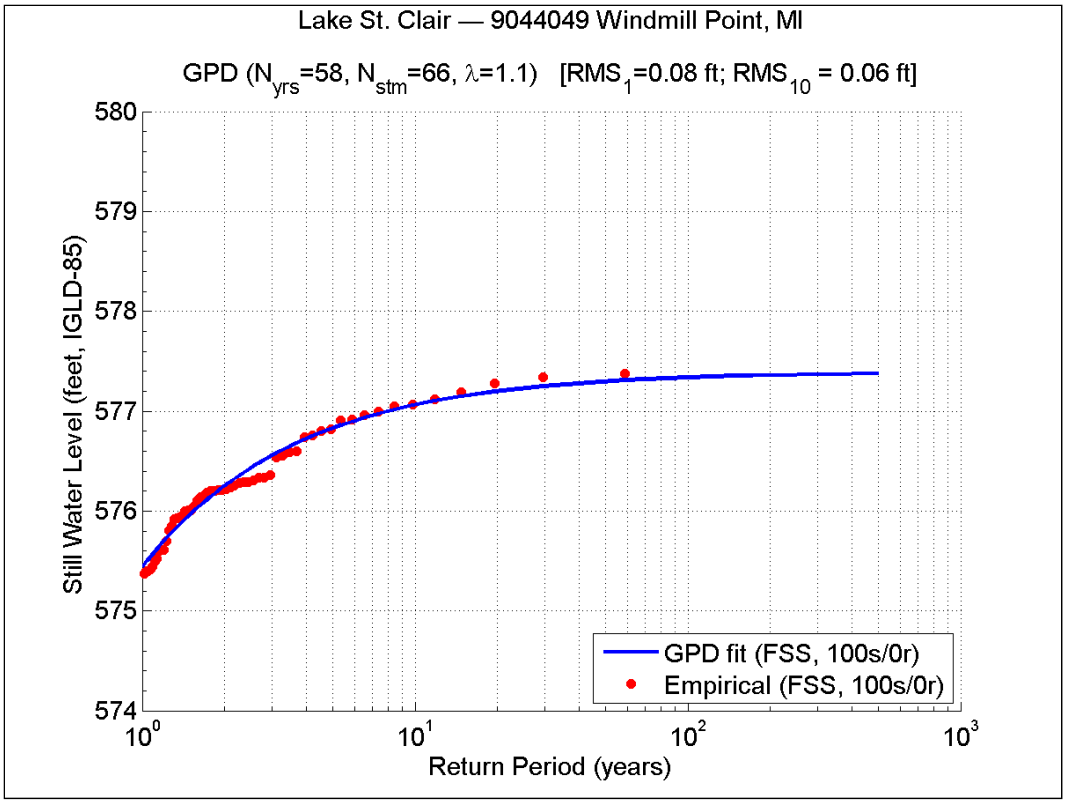



Figure D11. SWL RP plot from FSS (with convective storms) for Windmill Point, MI.

variability. For example, for St. Clair Shores the difference between the events ranked 100 and 150 was just 0.05 ft. Likewise, for Windmill Point this difference was 0.06 ft. The homogeneity among high-frequency events made the storm selection process more difficult for Lake St. Clair, but also made the final BFEs rather insensitive to the amount of sampled events.

The following are the SWL-CSS Q-Q, CDF and RP plots corresponding to the St. Clair Shores gage (Figures D12 through D14) and to the Windmill Point gage (Figures D15 through D17). Note that all these SWL-CSS plots include the October 1986 storms. Also, although the 144-storm CSS covers the entire 50 year period between 1960 and 2009, water level data for station 9034052 is only available from 1968 on, so just like the SWL-FSS the SWL-CSS is limited to 42 years (1968-2009). SWL-CSS for the Windmill Point gage accounts for the entire 50-year period between 1960 and 2009, as expected.

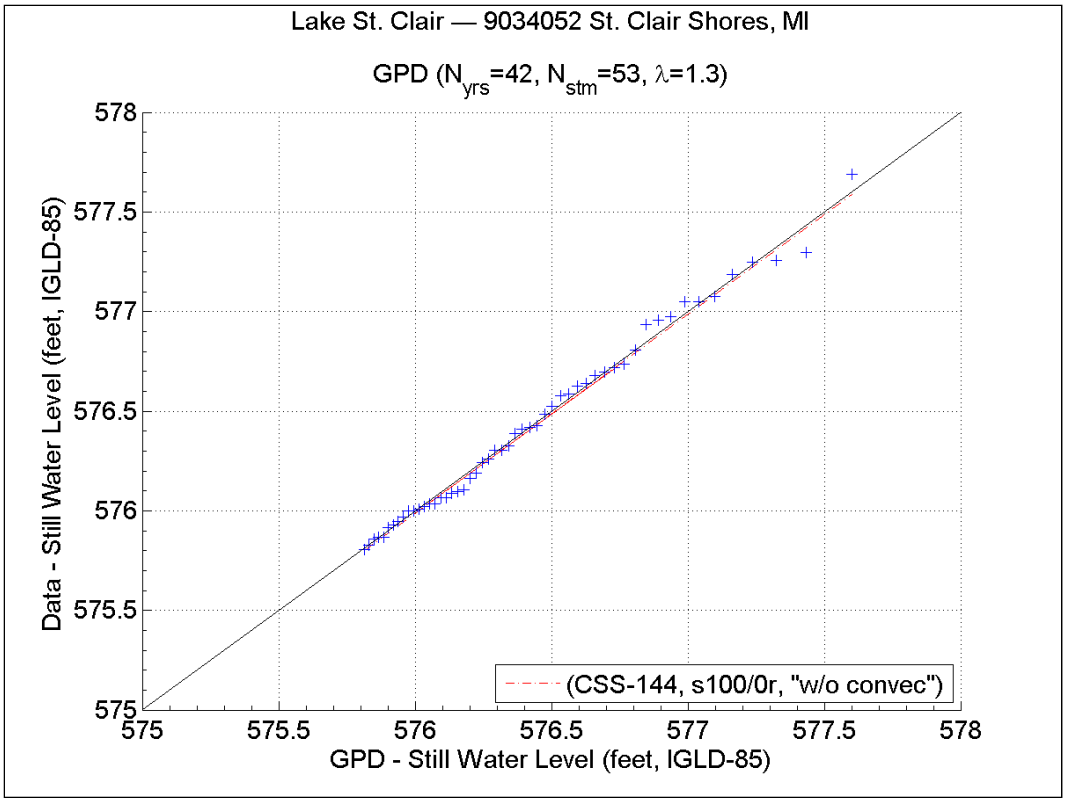



Figure D12. SWL Q-Q plot from CSS (100s/0r, no convective storms) for St. Clair Shores, MI.

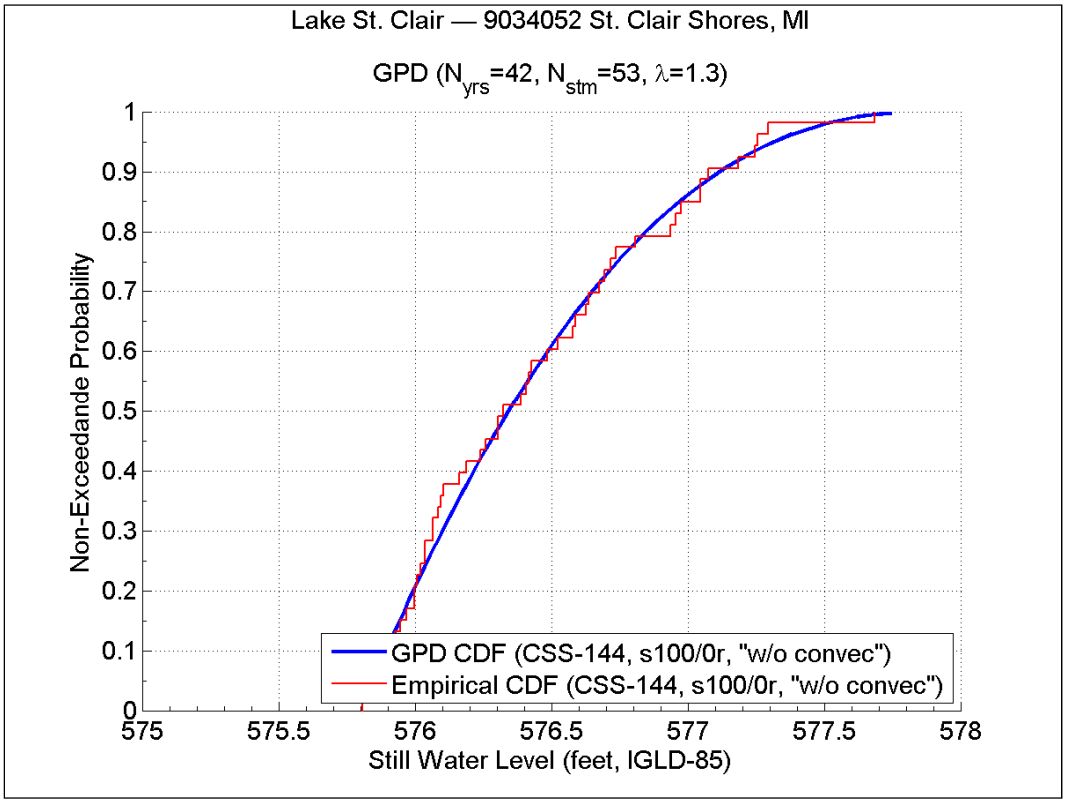



Figure D13. SWL CDF plot from CSS (100s/0r, no convective storms) for St. Clair Shores, MI.

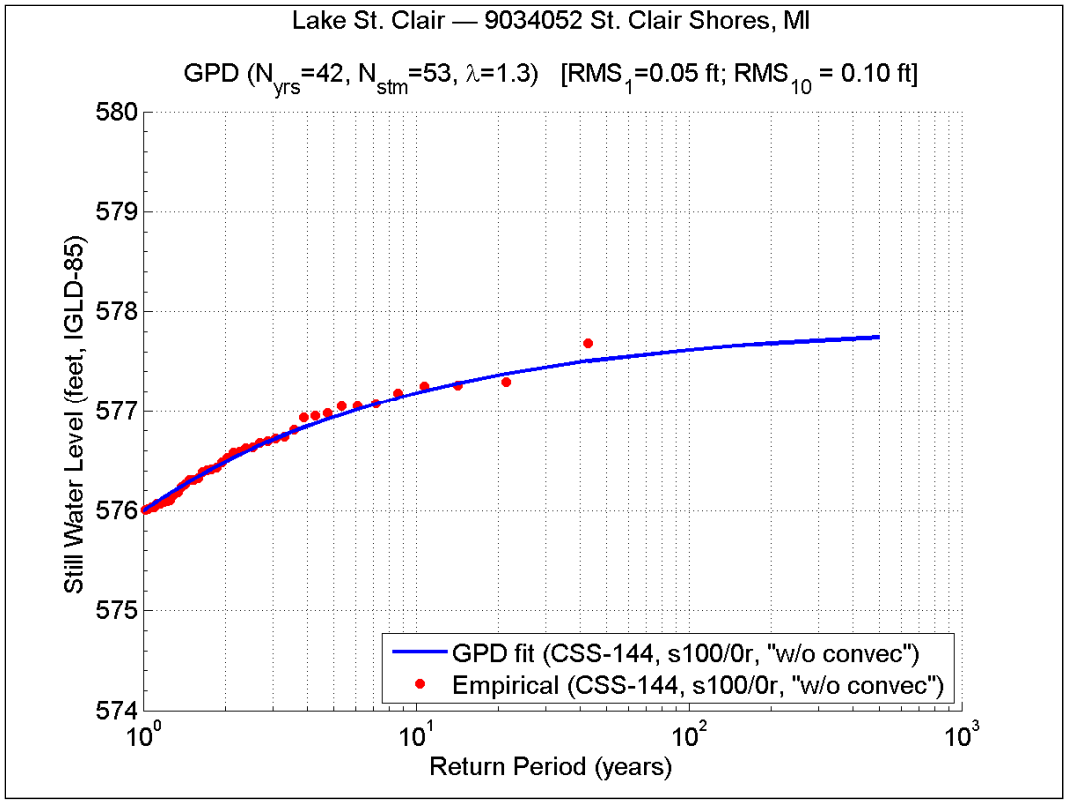



Figure D14. SWL RP plot from CSS (100s/0r, no convective storms) for St. Clair Shores, MI.

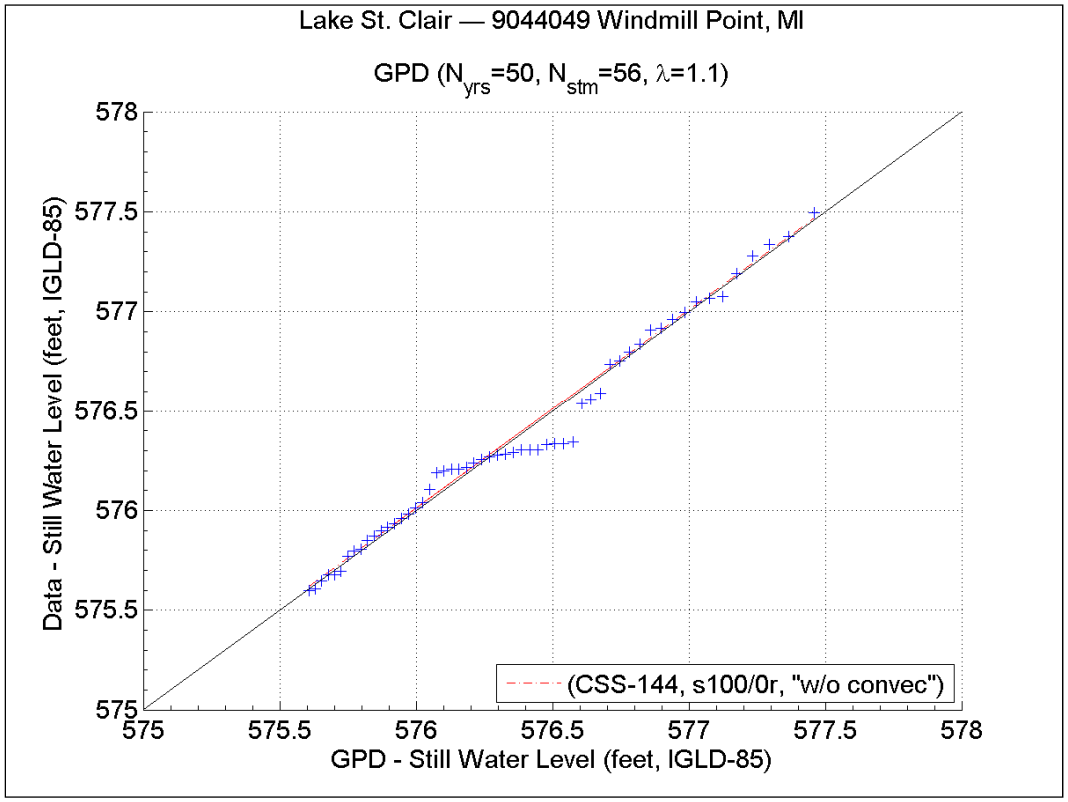



Figure D15. SWL Q-Q plot from CSS (100s/0r, no convective storms) for Windmill Point, MI.

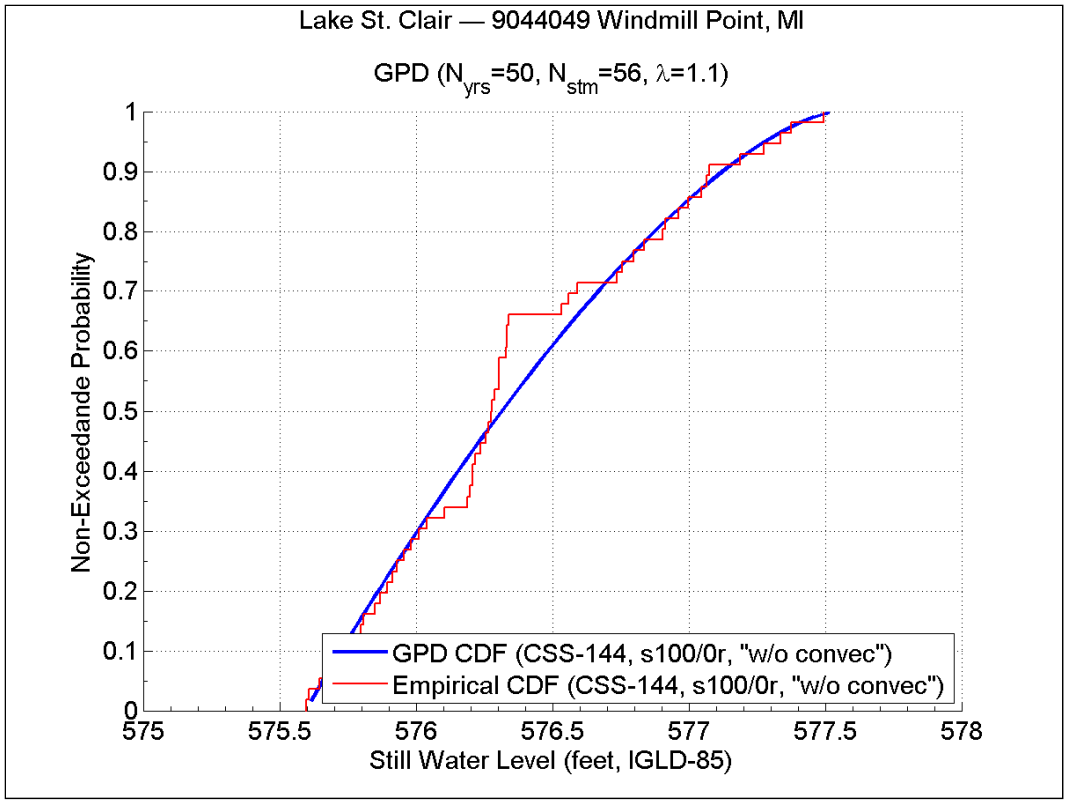



Figure D16. SWL CFD plot from CSS (100s/0r, no convective storms) for Windmill Point, MI.

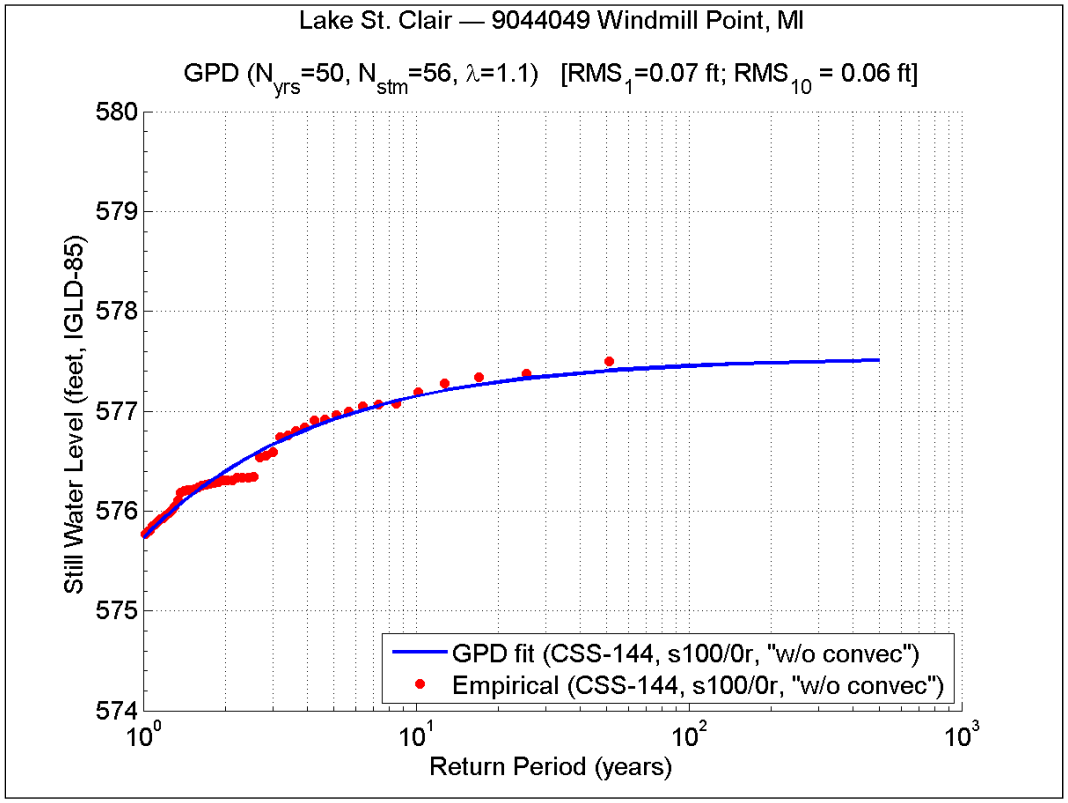



Figure D18. SWL RP plot from CSS (100s/0r, no convective storms) for Windmill Point, MI.

## Still water level: full storm set vs. composite storm set

Table D5 shows the SWL-FSS estimates corresponding to one percent and 0.2 percent annual chance values for the St. Clair Shores and Windmill Point gages, respectively. The SWL-FSS for both gages include summer convective events. As was the case for Lake Michigan, the SWL-FSS represents the "true" SWL distribution.

The SWL-CSS estimates corresponding to one percent and 0.2 percent annual chance values are shown on Table D6. Convective events are excluded from the SWL-CSS. To illustrate the changes in SWL-CSS due to the inclusion of the October 1986 event, which represents the highest overall water level at St. Clair Shores, the SWL values in Table D6.a were computed without this storm, while the values in Table D6.b were computed with it.

Table D5. SWL-CSS for one percent and 0.2 percent annual chance events.

| SWL-FSS without Oct/1986 event {100% surge - 0% wave} (ft, IGLD-85) | | |
|---|---|---|
| Probability (%) | 9034052 | 9044049 |
| 0.2 | 577.34 | 577.38 |
| 1 | 577.32 | 577.34 |

Table D6. SWL-CSS for one percent and 0.2 percent annual chance events.

| SWL-CSS without October 1986 event {100% surge - 0% wave} (ft, IGLD-85) | | |
|---|---|---|
| Probability (%) | 9034052 | 9044049 |
| 0.2 | 577.31 | 577.39 |
| 1 | 577.28 | 577.35 |
| SWL-CSS with October 1986 event {100% surge - 0% wave} (ft, IGLD-85) | | |
| Probability (%) | 9034052 | 9044049 |
| 0.2 | 577.75 | 577.51 |
| 1 | 577.62 | 577.46 |

Similarly to Lake Michigan, by comparing SWL-CSS (Table D6.a) and SWL-FSS (Table D5), it can be concluded that the effect of rejecting and replacing convective storms is minimal.

The differences between SWL-CSS (with and without the October 1986 storms) and SWL-FSS are shown on Table D7.a.

Table D7. SWL-CSS vs. SWL-FSS values for one percent and 0.2 percent annual chance events.

| [(SWL-CSS without October 1986 storm) – (SWL-FSS)] (ft) | | |
|---|---|---|
| Probability (%) | 9034052 | 9044049 |
| 0.2 | -0.04 | 0.01 |
| 1 | -0.04 | 0.01 |
| [(SWL-CSS with October 1986 storms) – (SWL-FSS)] (ft) | | |
| Probability (%) | 9034052 | 9044049 |
| 0.2 | 0.40 | 0.13 |
| 1 | 0.31 | 0.12 |

For St. Clair Shores, the differences between SWL-CSS and SWL-FSS are just 0.04 feet for both the one percent and 0.2 percent annual chance events. Likewise, for Windmill Point these differences are a mere 0.01 feet.

Table D7.b shows the differences between the SWL-CSS including the October 1986 event and the SWL-FSS. For St. Clair Shores, the SWL corresponding to one percent and 0.2 percent annual chance increased by 0.31 feet and 0.40 feet, respectively. These differences show that it is imperative to include events that occur near the highest overall water level, even if the event itself is below the POT threshold.

For Lake Michigan, all of the highest water levels were captured by events sampled through the POT approach, but for Lake St. Clair, adjustments were required and it was necessary to revisit the list of events and add the October 1986 event post-POT.

## Lake St. Clair 144-storm CSS

The following list (in chronological order) constitutes the final 144 storms selected as the CSS for Lake St. Clair Flood Hazard Mapping production. Note that storms 11, 16, 81, 98, 117 and 121 were rejected from the final set, reducing the number of storms from 150 to 144.

| Storm # | Peak Date (UTC) | | | |
|---|---|---|---|---|
| | Year | Month | Day | Hour |
| 1 | 1986 | 10 | 6 | 0 |
| 2 | 1960 | 11 | 16 | 12 |
| 3 | 1961 | 3 | 9 | 20 |
| 4 | 1961 | 4 | 16 | 17 |
| 5 | 1962 | 12 | 29 | 20 |
| 6 | 1963 | 3 | 17 | 15 |
| 7 | 1963 | 4 | 4 | 6 |
| 8 | 1963 | 10 | 3 | 4 |
| 9 | 1969 | 5 | 8 | 18 |
| 10 | 1964 | 3 | 5 | 18 |
| 11 | | | | |
| 12 | 1977 | 5 | 8 | 21 |
| 13 | 1990 | 2 | 24 | 16 |
| 14 | 1966 | 1 | 3 | 17 |
| 15 | 1966 | 3 | 24 | 2 |
| 16 | | | | |
| 17 | 1968 | 5 | 16 | 18 |
| 18 | 1968 | 11 | 20 | 0 |
| 19 | 1968 | 12 | 5 | 18 |
| 20 | 1978 | 11 | 27 | 10 |
| 21 | 2003 | 4 | 17 | 8 |
| 22 | 1970 | 3 | 26 | 14 |
| 23 | 1970 | 4 | 2 | 11 |
| 24 | 1970 | 4 | 19 | 20 |
| 25 | 1970 | 11 | 20 | 11 |
| 26 | 2004 | 3 | 16 | 20 |
| 27 | 1971 | 3 | 19 | 12 |
| 28 | 1971 | 12 | 30 | 14 |
| 29 | 1972 | 4 | 22 | 8 |
| 30 | 1972 | 11 | 14 | 10 |
| 31 | 1973 | 1 | 4 | 3 |
| 32 | 1973 | 2 | 6 | 11 |

| Storm # | Peak Date (UTC) | | | |
|---|---|---|---|---|
| | Year | Month | Day | Hour |
| 33 | 1973 | 3 | 17 | 14 |
| 34 | 1973 | 4 | 9 | 20 |
| 35 | 1973 | 10 | 28 | 6 |
| 36 | 1974 | 3 | 12 | 20 |
| 37 | 1974 | 4 | 8 | 14 |
| 38 | 1974 | 12 | 2 | 9 |
| 39 | 1975 | 3 | 28 | 7 |
| 40 | 1975 | 8 | 31 | 7 |
| 41 | 1975 | 10 | 18 | 9 |
| 42 | 1990 | 5 | 4 | 21 |
| 43 | 1976 | 3 | 5 | 12 |
| 44 | 1976 | 3 | 12 | 7 |
| 45 | 1976 | 4 | 25 | 17 |
| 46 | 1977 | 3 | 18 | 12 |
| 47 | 1977 | 4 | 4 | 23 |
| 48 | 1977 | 4 | 27 | 0 |
| 49 | 1977 | 12 | 5 | 20 |
| 50 | 1977 | 12 | 20 | 15 |
| 51 | 1975 | 4 | 3 | 0 |
| 52 | 1999 | 4 | 23 | 19 |
| 53 | 1978 | 4 | 6 | 19 |
| 54 | 1979 | 4 | 12 | 11 |
| 55 | 1979 | 4 | 17 | 9 |
| 56 | 1979 | 12 | 27 | 22 |
| 57 | 1980 | 4 | 14 | 14 |
| 58 | 1980 | 11 | 27 | 18 |
| 59 | 1987 | 4 | 4 | 22 |
| 60 | 1981 | 12 | 1 | 9 |
| 61 | 1982 | 4 | 6 | 8 |
| 62 | 1982 | 10 | 10 | 7 |
| 63 | 1983 | 4 | 9 | 18 |
| 64 | 1983 | 4 | 17 | 8 |

| Storm # | Peak Date (UTC) | | | |
|---|---|---|---|---|
| | Year | Month | Day | Hour |
| 65 | 1985 | 11 | 28 | 10 |
| 66 | 1984 | 2 | 28 | 13 |
| 67 | 1984 | 3 | 29 | 2 |
| 68 | 1984 | 4 | 5 | 1 |
| 69 | 1984 | 5 | 28 | 22 |
| 70 | 1985 | 1 | 3 | 18 |
| 71 | 1983 | 11 | 28 | 11 |
| 72 | 1984 | 5 | 3 | 21 |
| 73 | 1985 | 3 | 31 | 12 |
| 74 | 1985 | 11 | 16 | 14 |
| 75 | 1985 | 11 | 22 | 12 |
| 76 | 1986 | 12 | 2 | 6 |
| 77 | 1986 | 12 | 8 | 22 |
| 78 | 1978 | 10 | 16 | 13 |
| 79 | 1987 | 3 | 10 | 3 |
| 80 | 1987 | 12 | 15 | 15 |
| 81 | | | | |
| 82 | 1989 | 2 | 17 | 0 |
| 83 | 1989 | 10 | 19 | 15 |
| 84 | 1975 | 3 | 14 | 20 |
| 85 | 1990 | 11 | 5 | 20 |
| 86 | 1990 | 12 | 3 | 17 |
| 87 | 1990 | 12 | 31 | 14 |
| 88 | 2005 | 1 | 6 | 7 |
| 89 | 1991 | 10 | 29 | 5 |
| 90 | 1991 | 12 | 3 | 5 |
| 91 | 1992 | 11 | 2 | 9 |
| 92 | 1993 | 1 | 10 | 16 |
| 93 | 1993 | 9 | 26 | 4 |
| 94 | 2000 | 5 | 19 | 14 |
| 95 | 1994 | 11 | 27 | 18 |
| 96 | 1995 | 2 | 27 | 8 |

| Storm # | Peak Date (UTC) | | | |
|---|---|---|---|---|
| | Year | Month | Day | Hour |
| 97 | 1995 | 10 | 6 | 3 |
| 98 | | | | |
| 99 | 1970 | 12 | 11 | 17 |
| 100 | 1996 | 4 | 4 | 19 |
| 101 | 1996 | 4 | 30 | 2 |
| 102 | 1995 | 12 | 19 | 23 |
| 103 | 1997 | 3 | 4 | 3 |
| 104 | 1997 | 3 | 14 | 9 |
| 105 | 1997 | 10 | 27 | 5 |
| 106 | 1997 | 11 | 14 | 11 |
| 107 | 1998 | 1 | 9 | 5 |
| 108 | 1998 | 2 | 18 | 2 |
| 109 | 1998 | 4 | 9 | 22 |
| 110 | 1998 | 3 | 21 | 12 |
| 111 | 1978 | 3 | 25 | 18 |
| 112 | 1999 | 3 | 6 | 15 |
| 113 | 1999 | 4 | 9 | 13 |
| 114 | 1999 | 12 | 14 | 18 |
| 115 | 2000 | 2 | 29 | 23 |
| 116 | 2000 | 4 | 20 | 20 |
| 117 | | | | |
| 118 | 2001 | 2 | 15 | 1 |
| 119 | 2001 | 10 | 16 | 14 |
| 120 | 2002 | 1 | 31 | 14 |
| 121 | | | | |
| 122 | 1993 | 4 | 1 | 9 |
| 123 | 2003 | 4 | 7 | 20 |
| 124 | 2003 | 5 | 5 | 13 |
| 125 | 2003 | 12 | 5 | 16 |
| 126 | 2004 | 1 | 5 | 6 |
| 127 | 1993 | 5 | 12 | 21 |
| 128 | 2004 | 5 | 31 | 2 |

| Storm # | Peak Date (UTC) | | | |
|---|---|---|---|---|
| | Year | Month | Day | Hour |
| 129 | 2004 | 11 | 24 | 21 |
| 130 | 2005 | 12 | 25 | 22 |
| 131 | 2006 | 3 | 13 | 8 |
| 132 | 2006 | 3 | 17 | 9 |
| 133 | 2006 | 12 | 1 | 11 |
| 134 | 2007 | 3 | 29 | 11 |
| 135 | 2007 | 4 | 12 | 1 |
| 136 | 2008 | 1 | 11 | 1 |
| 137 | 1972 | 4 | 7 | 14 |
| 138 | 2008 | 5 | 11 | 19 |
| 139 | 2008 | 12 | 19 | 14 |
| 140 | 2008 | 12 | 31 | 5 |
| 141 | 1992 | 1 | 16 | 0 |
| 142 | 2009 | 3 | 12 | 18 |
| 143 | 2009 | 10 | 23 | 17 |
| 144 | 2009 | 12 | 9 | 9 |
| 145 | 2009 | 12 | 25 | 16 |
| 146 | 1975 | 4 | 3 | 0 |
| 147 | 1978 | 11 | 27 | 10 |
| 148 | 1999 | 4 | 23 | 19 |
| 149 | 2003 | 4 | 17 | 8 |
| 150 | 2004 | 3 | 16 | 20 |

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From - To)* |
|---|---|---|
| September 2012 | Final | |

**4. TITLE AND SUBTITLE**

Statistical Analysis and Storm Sampling Approach for Lakes Michigan and St. Clair

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

Norberto C. Nadal-Caraballo, Jeffrey A. Melby and Bruce A. Ebersole

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Coastal and Hydraulics Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-619

**8. PERFORMING ORGANIZATION REPORT NUMBER**

ERDC/CHL TR-12-19

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

U.S. Army Corps of Engineers
Detroit District
477 Michigan Ave.
Detroit, MI 48226

FEMA Region V
536 S. Clark St., 6th Floor
Chicago, IL 60605

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for public release; distribution is unlimited.

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

A new methodology is proposed for Great Lakes flood hazard mapping. The methodology includes a process for sampling and screening storm events and computing water level probabilities based on high-fidelity modeling of significant storm events. A technical analysis framework is provided to construct accurate extremal distributions of total water levels and to accurately estimate base flood elevations. High-resolution, high-fidelity modeling of all historical storm events is simply not feasible due to time, computational and funding constraints. Therefore, the recommended approach is to screen and sample historical events to select the minimum number of events required to accurately model the total water level extremal distributions. This study focused on evaluating the validity of the recommended statistical analysis and storm sampling approach, and determining the adequate storm sample size. For this purpose, several tasks were performed, including: computation of storm-surrogate waves and water levels; definition of full sample and composite storm sets; and evaluation of the statistical analysis approach for a record length of 50 years.

**15. SUBJECT TERMS**

Extremal distributions
Flood hazard
Great Lakes
Lake Michigan
Statistical analysis
Storm sampling
Water levels
waves

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON: Ty Wamsley** |
|---|---|---|---|---|---|
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | | | **19b. TELEPHONE NUMBER** *(include area code)* |
| Unclassified | Unclassified | Unclassified | Unclassified | 155 | (601) 634-2099 |